FUNCTIONS OF THE SEVENTEEN KILODALTON PROTEIN IN ESCHERICHIA COLI PROTEOSTASIS

by
Clifford Wade Sandlin

A dissertation submitted to Johns Hopkins University in conformity with the requirements for the degree of Doctor of Philosophy

Baltimore, Maryland
July, 2016

Abstract

Outer membrane (OM) biogenesis is critical for the survival of Gram-negative bacteria. This process depends on periplasmic protein homeostasis, or proteostasis. Proteostasis depends on quality control (QC) networks. However, the periplasmic network is dense with aggregation-prone uOMPs and lacks ATP for disaggregases. To combat this, Seventeen Kilodalton Protein (Skp) is believed to provide chaperone activity (CA) and antiaggregation activity (AA).

The Skp CA is proposed to work by enclosing uOMPs within a cavity of trimeric Skp. This cavity is smaller than the dimensions of free uOMPs. To resolve this paradox, I used SANS to measure the R_g of apo- and holo-Skp. SANS and modeling suggest that Skp adapts differently to uOmpW and uOmpA.

Next, I revisited the evidence for CA and sought to verify Skp CA with OMPs. SDS-PAGE delayed-folding assays with uOmpW, uOmpLa and uBamA only showed Skp CA with uOmpW. This suggested that Skp CA does not apply to all uOMPs.

In addition, I asked whether Skp is always a trimer. AUC sedimentation equilibrium (SE) reveals that some Skp is monomeric at physiological concentrations. This link between Skp trimerization and concentration led to a “switching hypothesis”, that functional Skp trimers are upregulated by stress.

In support of switching, sedimentation velocity (SV) shows a monomeric variant of Skp lacks AA. Skp AA and CA varied between uOmpA, uOmpLa, and uBamA-TM. Surprisingly, I found that AA did not depend on uOMP structure or correlate with CA.

Finally, with fluorescence and SV I show that free uOmpLa and uBamA aggregate at ~10 nM in 150 mM NaCl, 80 mM urea, and pH 8.0. This suggests that the nanomolar affinity of Skp is well-suited to prevent uOMP aggregation in cells.

I conclude that Skp is mainly a quality control antiaggregation factor, and has only a minor function as an OMP chaperone. I propose a shift away from the chaperone paradigm for Skp, and suggest that Skp evolved to serve as a "buffer", for uOMPs, absorbing excess uOMPs when periplasmic conditions do not favor folding.

Thesis Advisor:	Dr. Karen Fleming
Second Reader:	Dr. Doug Barrick
Thesis Committee:	Dr. Richard Cone
	Dr. Bertrand Garcia-Moreno E.
	Dr. Julette Lecomte

INTENDED TO BE BLANK

For all that I have left behind

And all that I have left to find

Acknowledgements

This document, such as it is, is a culmination of nine years of my life. I mostly worked and sacrificed alone reach this point. However, others contributed much to the effort. I will do my best to acknowledge them in no particular order, save one.

My mother Rumi provided the most important support during the most difficult and trying moments I faced on this journey. She provided an eager ear on hundreds of occasions, and excellent advice on just as many. She has given me a number of material gifts along the way that have eased my route, especially in times of high stress. Whenever I was forced to deal the worst character traits in others, she displayed the grace, humility, strength, and sense of honor and justice that represent the best of all of us.

My father Robert instilled me with a love of science and technology at an early age. This early exposure was essential, I think, to fully motivate my decision to pursue the path of a nonconventional student. My phone conversations with him throughout graduate school also reinforced my skills in deductive logic.

I would like to offer special thanks to Dr. Richard Cone, who possesses a positive attitude and a love of science that are as infectious as his advocacy of the underemphasized models of molecular diffusion.

Dr. Karen Fleming, my Ph.D. advisor, provided me with excellent technical and professional advice on many occasions, and I could not have completed this work without her assistance. In terms of technical knowledge, she trained me in the use of the analytical ultracentrifuge, and in best practices for scientific communication. Her most important influences on my development as a scientist are her practice of focusing on

critical questions, her tendency to carefully weigh risk, and her gift for setting realistic goals. Early on, she forced me to let go of a difficult project, on the Sam50 protein, that was holding me back. This lesson of learning how to let go of unproductive work was very difficult for me personally, yet it was absolutely necessary for me to achieve success in a world with practical limitations.

Dr. Nathan Zaccai trained me in protein purification on my arrival at the Fleming lab, and introduced me to the technique of neutron scattering. He taught me a sense of urgency, the importance of being first in science, and when to make discovery the top priority. He taught me how to "roll with the punches" and acted with courtesy even when faced with the opposite. He taught me professionalism and inspired me to take risks.

My lab mates through the years, as some of the few people who really understand the unique challenges I faced, have been an important support structure. In particular, graduate students Sarah McDonald and Dagan Marx have been excellent sounding boards for ideas and have been quick to raise my spirits when needed.

A number of teachers from my undergraduate education and internships were a tremendous influence on me intellectually. Dr. Kai-Sheng Song and Dr. John Quintanilla taught me a great deal about how to think about data using statistics, and I tried to bring this perspective to my work whenever it was possible. Dr. Neal Brand introduced me to formal deductive logic and Dr. William Cherry enrolled me in the mathematics degree. The latter has been critical to my success as a graduate student and scientist and has fundamentally changed me for the better. Dr. Lloyd Fitzpatrick taught me to think about organisms as members of communities, which is a perspective sometimes lost to

laboratory scientists. My undergraduate mentor, Dr. Rebecca Dickstein, introduced me to biochemical research and taught me how to think about molecular events in cells.

Additional special thanks are due for the contributions of thesis committee members Dr. Doug Barrick and Dr. Juliette Lecomte. The original draft of my thesis was not well-structured for its intended audience, and chapters 1, 3 and 5 had to be rewritten. During this rewriting process, Doug and Juliette made a special effort to help me reach my graduation deadline that went far beyond the requirements of their job descriptions. Dr. Barrick also provided extremely useful feedback on the description of the chaperone activities. Dr. Lecomte has an eye for detail that is truly shocking to the uninitiated.

Also, thanks to graduate students Peregrine Bell-Upp, Scott Nichols, Mollie Rappe, Jeremy Anderson, Chris King, and Shannen Cravens. You were all humble, down-to-earth, sincere people. Also thanks to Sammy the Black Cat for being one of my best friends on my long walks home alone at night and to all present and past staff of Miss Shirley's Café at Roland Park for giving me a grounding point and public man-cave. Finally, thank you Miranda for teaching me that there is more in this life than reason.

Abbreviations

AA	Antiaggregation Activity
AFM	Atomic Force Microscopy
AR	Aggregation Rate
Arg	Arginine
ATP	Adenosine Triphosphate
AUC	Analytical Ultracentrifugation
CA	Chaperone Activity
CD	Circular Dichroism
$diC_{10}PC$	1,2-didecanoyl-*sn*-glycero-3-phosphocholine
$diC_{11}PC$	1,2-diundecanoyl-*sn*-glycero-3-phosphocholine
D_{CM}	Distance between the centers of mass of two objects.
$g(S_{20,W}*)$	The absorbance distribution of species in sedimentation velocity as a function of their apparent sedimentation coefficients, adjusted to standard conditions.
IPTG	isopropyl β-D-1-thiogalactopyranoside
LB	Loading Buffer
LUV	Large Unilamellar Vesicle
mpd	Minutes post-dilution. The time after rapid dilution of some stock including protein (typically Skp), into a solution that starts a reaction.
M_W	Molecular Weight
OM	Outer Membrane

OMP	Outer Membrane Protein
PC	Phosphatidylcholine
PG	phophatidylglycerol
pH	-log[Hydrogen Ion Concentration], M
Phe	Phenylalanine
PP	Periplasmic domain
R_g	Radius of Gyration
R_H	Hydrodynamic radius
S*	Apparent sedimentation coefficient
$S_{20,W}$*	Sedimentation coefficient converted to standard conditions of 20 °C in water
SANS	Small Angle Neutron Scattering
SD	Standard Deviation
SDS	Sodium Dodecyl Sulfate
SDS-PAGE	Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis
SEM	Standard error of the mean
TCEP	tris(2-carboxyethyl) phosphine hydrochloride
TCEP-HCl	TCEP
TM	Transmembrane
Tris-HCl	Tris(hydroxymethyl) aminomethane Hydrochloride
Trp	Tryptophan
Tyr	Tyrosine
uOMP	Unfolded Outer Membrane Protein

List of Proteins

BamA	β-barrel assembly machinery subunit A
BamA-TM	β-barrel assembly machinery subunit A, mature, β-barrel domain residues 399-790
DegP	Periplasmic serine endoprotease DegP
FAI	Full Alanine Interface Skp with all trimerization interface residues altered to alanine
FkpA	FKBP-type petidyl-prolyl cis-trans isomerase
OmpA	Outer membrane protein A, mature
OmpA-TM	Outer membrane protein A, mature, β-barrel domain residues 1-171
OmpLa	Outer membrane protein phospholipase A, mature
OmpW	Outer membrane protein W, mature
PDI	Protein disulfide isomerase
PPI	Peptidyl-prolyl cis-trans isomerase
TM	Transmembrane domain. e.g., OmpA-TM and BamA-TM
SAI	Skp Y124A
Skp	When the text refers to a situation *in vivo*, the mature seventeen kilodalton protein. In all *in vitro* experiments, Skp refers to the mature seventeen kilodalton protein with the signal sequence removed a histidine affinity purification tag added to the C-terminus.
Skp-His	Skp
SurA	Survival protein A

TEV	Tobacco etch virus protease
uBamA	Unfolded β-barrel assembly machinery subunit A
uBamA-TM	Unfolded β-barrel assembly machinery subunit A, mature, β-barrel domain
uOmpA	Unfolded outer membrane protein A, mature
uOmpA-TM	Unfolded outer membrane protein A, mature, β-barrel domain residues 1-171
uOmpLa	Unfolded outer membrane protein phospholipase A, mature
uOmpW	Unfolded outer membrane protein W, mature
uOmpX	Unfolded outer membrane protein X
WT Skp	Seventeen kilodalton protein with an N-terminal Histidine affinity tag that has been removed by TEV.

Table Of Contents

List of Tables

List of Figures

Chapter 1
Introduction

1.1 The Importance of Outer Membrane Proteostasis and the Seventeen Kilodalton Protein (Skp)

Total prokaryote biomass approaches 60-100% of the biomass of plant life (Whitman, Coleman et al., 1998). A large amount of this population is Gram-negative bacteria. These Gram-negatives engage in commensal (Alekshun and Levy, 2006) and pathogenic (Koli, Sudan et al., 2011) symbioses with humans. Therefore, an increased understanding of the surfaces, or envelopes of these bacteria will inform our understanding of basic biology and medicine.

Gram-negative envelopes consist of an inner membrane (IM), which contains mostly α-helical proteins, and an outer membrane (OM), which contains proteins of β-barrel architecture (figure 1.1) (Nakamura and Mizushima, 1976, Vogt and Schulz, 1999, Schulz, 2002, Hiller and Wagner, 2009, Bratanov, Balandin et al., 2015). These outer membrane proteins (OMPs) make up 67% wt/wt of the *Escherichia coli* OM and cover most its surface (Jaroslawski, Duquesne et al., 2009). Such abundance should suggest that OMPs are essential for *E. coli* to survive. Consistent with this idea, disrupting OMP steady-state expression, also called biogenesis, results in a lethal phenotype (Voulhoux, Bos et al., 2003). The biogenesis of OMPs is therefore an important research topic.

OMP biogenesis requires control over synthesis, transport, folding, and degradation (figure 1.2). This tight regulation is referred to as proteostasis (Balch, Morimoto et al., 2008). Proteostasis of OMPs presents a special problem for biology,

because OMPs are highly hydrophobic (Moon, 2011). Hydrophobicity of OMPs makes them prone to self-associate, or aggregate (Danoff and Fleming, 2011, Danoff and Fleming, 2015). This aggregation correlates with a drop in folding efficiency into membranes (figure 1.3) (Burgess, Dao et al., Ebie Tan, Burgess et al., 2010). Furthermore, if OMPs aggregate *in vitro* they can form structures called amyloid fibrils, which could be harmful *in vivo* (figure 1.4) (Bednarska, Schymkowitz et al., 2013, Danoff and Fleming, 2015, Tuttle, Comellas et al., 2016).

uOMP aggregation is prevented by proteins, called "quality control" factors (Kubota, 2009). Quality control (QC) describes many functions that work together to promote proteostasis. Some of these QC functions are degradation, modification, and transport. For example, DegP degrades OMPs, DsbA modifies disulfide bonds, and SecA transports OMPs across the inner membrane (IM) (Hansen and Hilgenfeld, 2013, Denks, Vogt et al., 2014, Lu and Holmgren, 2014).

Other QC proteins prevent OMP aggregation and promote folding (figure 1.2). These proteins include BamA, SurA, PPI-D, FkpA, Spy, HdeA, HdeB, and the Seventeen Kilodalton Protein (Skp) (Missiakas, Betton et al., 1996, Tapley, Franzmann et al., 2010, Webb, Heinz et al., 2012, Ding, Yang et al., 2015, Stull, Koldewey et al., 2016). These are referred to as the OM biogenesis proteins (Fleming, 2015, Kleinschmidt, 2015, Rollauer, Sooreshjani et al., 2015). BamA, is part of a complex, BamABCDE that is essential for OMP folding (Voulhoux, Bos et al., 2003, Gu, Li et al., 2016). SurA, is essential for OMP expression and accelerates OMP folding *in vitro* (Rizzitello, Harper et al., 2001, Thoma, Burmann et al., 2015). PPI-D sits near the mouth the Sec translocon and, with SurA may provide a route for uOMPs to reach BamABCDE (Sachelaru,

Petriman et al., 2014). FkpA, Spy, and the HdeA/B promote folding in response to environmental stresses. FkpA is activated by heat (Ge, Lyu et al., 2014), Spy by denaturants (Stull, Koldewey et al., 2016), and HdeA/B by acidic conditions (Yohannes, Barnhart et al., 2004). Thus, except for Skp, each OM biogenesis factor has been shown to play a specific role. The function of Skp is not obvious from its phenotype. Skp is nonessential in rich media (Schäfer, Beck et al., 1999, Rizzitello, Harper et al., 2001, Sklar, Wu et al., 2007). However, when SurA or DegP are deleted, Skp becomes essential (Sklar, Wu et al., 2007). This synthetic lethality, and the specialization seen in the other chaperones, led me to hypothesize that Skp has a functional specialty like FkpA, Spy, and HdeA/B. Such classification requires that I identify the QC activities Skp possesses. Therefore, I will next describe all the activities in terms of kinetics.

1.2 Types of Quality Control and Molecular Chaperone Activities

Skp could have a number of distinct biochemical "activities", that each may impact OMP biogenesis. As I define these activities, I use the term "client" to mean the target of an activity. Also, I refer to the putative chaperone as "the candidate". Some of these activities are organized into a subgroup called chaperone activities (CA).

Quality Control Activities

Antiaggregation Activity (AA)

AA means that binding the client competes kinetically with aggregation. In figure 1.5, the rate constant, k_{on}, of client binding is larger than k_{agg}, the rate of aggregation. Also, the client-candidate complex does not aggregate (figure 1.5, $k_{cdis} >> k_{cagg}$). A classic example is the protein trigger factor, which binds clients as they emerge from the

ribosome and before they can aggregate (Hoffmann, Bukau et al., 2010, Saio, Guan et al., 2014).

Binding Activity

The protein binds the client. This means $k_{on} > k_{off}$ (see figure 1.5). Binding activity may manifest as antiaggregation activity. However, if the rate constant $k_{cdis} << k_{cagg}$ (figure 1.5), aggregation still results. Examples are the small heat shock proteins (sHSPs), which have many modes of binding clients and each other (Haslbeck and Vierling, 2015).

Catabolism/Degradation

The protein acts proteolytically to hydrolyze the client. In *E. coli*, uOMPs are degraded by DegP (Hansen and Hilgenfeld, 2013).

Post-translational Modification

Chemical addition of functional groups and disulfide bond isomerization fall in this category. An example is the DsbA protein disulfide isomerase (PDI) (Lu and Holmgren, 2014).

Chaperone Activities

This special group of QC activities promotes folding or assembly. Some proteins can have multiple chaperone activities (CA). Refer to appendix I for an optional commentary about recent findings that may impact the way chaperones are defined.

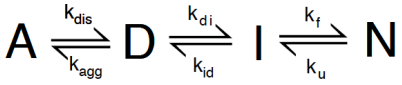
Foldase Activity

This CA is a catalytic increase in the folding rate due to interaction with the chaperone. In figure 1.5, the rate constant k_f is the target of foldase activity.

An example of a foldase chaperone is heat shock protein 70 (HSP70) (Young, Agashe et al., 2004).

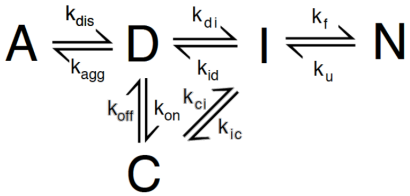
Holding Activity

Holding activity is easiest to understand as an extreme case of two alternative kinetic schemes. The first scheme is the chaperone-independent pathway. The second scheme includes the chaperone.



Scheme 1

Scheme 1 (see also figure 1.5, A ↔ D ↔ I ↔ N) illustrates folding through an intermediate. For holding activity to make sense, the constants in scheme 1 must fulfill many assumptions. First, folding (I ↔ N) must not be rate-limiting. This means the formation of I is slower than folding ($k_{di}<k_f$). In addition, favorable aggregation ($k_{agg}>k_{dis}$) must compete kinetically ($k_{agg}>k_{di}$) with favorable membrane binding ($k_{di}>k_{id}$). These assumptions result in a faster rate of aggregation (A ← D) than intermediate formation (D → I) that slows the formation of N.



Scheme 2

Scheme 2 (see also figure 1.5, A ↔ D ↔ C ↔ I ↔ N, ignore A_c) is similar to scheme 1, but adds another intermediate, C. Suppose the assumptions of scheme 1 are

true for scheme 2. Holding activity results when the constants k_{on} and k_{ci} that favor the intermediate, I, are both larger than all the constants k_{agg}, k_{off}, and k_{ic}. To summarize, holding activity means that k_{on} and k_{ci} exceed all other rate constants in scheme 2, resulting in faster formation of N than for scheme 1.

An important example of a protein with holding activity is the GroEL/GroES chaperonin, which also has foldase activity (Goloubinoff, Gatenby et al., 1989, Gupta, Haldar et al., 2014, Ishimoto, Fujiwara et al., 2014). The original chaperone, nucleoplasmin, has holding activity with histones (Laskey, Honda et al., 1978).

Disaggregase Activity

Disaggregases are chaperones that increase the rate constant (k_{dis} in figure 1.5) at which aggregates dissociate.

An example is the ATP-dependent chaperone HSP104 (Aguado, Fernandez-Higuero et al., 2015).

Assembly Activity

This oft-forgotten activity (Ellis, 2013) accelerates the rate of complex formation (k_a in figure 1.5).

Skp is thought to possess three of these activities: binding, AA, and holding activity. The evidence that Skp has these properties is discussed below.

1.3 Current Views on the Role of Skp in the Periplasm Lead to Three Questions

Skp is often called a chaperone in the literature. The conclusion that Skp is a chaperone is mostly based on results from genetic knockouts *in vivo* and studies of binding *in vitro*. *In vivo,* knockout of Skp activates a stress response in *E. coli* (Missiakas, Betton et al., 1996). This stress response, called the sigmaE response, is named for the sigmaE transcription factor (Kim, 2015). SigmaE upregulates many genes, including Skp (Rhodius, Suh et al., 2006). Importantly, sigmaE is activated by unfolded OMPs in the periplasm (Ades, 2004). This connection between Skp and sigmaE led to the hypothesis that Skp is a chaperone for OMP folding (Missiakas, Betton et al., 1996). Consistent with this notion, Skp deletion reduces uOMP solubility in the periplasm (Schäfer, Beck et al., 1999). Also, Skp interacts genetically with DegP, which degrades uOMPs (Schäfer, Beck et al., 1999). Moreover, Skp binds uOMPs, with nanomolar affinity (Qu, Mayer et al., 2007, Moon, Zaccai et al., 2013).

Structural evidence also supports the model that Skp is a QC protein. The Skp quaternary structure (figure 1.6) is similar to prefoldin, which delivers cytoskeletal proteins to chaperonin II (Korndörfer, Dommel et al., 2004, Stirling, Bakhoum et al., 2006, Millan-Zambrano and Chavez, 2014). Furthermore, NMR studies revealed that uOMPs bound within the Skp trimer cavity are held as unstructured and dynamic "fluid globules" (Burmann, Wang et al., 2013). However, these NMR data suggest a question: *Do free uOMPs contract to fit inside this cavity, or does Skp adapt to them* (figure 1.7)?

If uOMPs must collapse to enter Skp, then such compaction might slow Skp-uOMP binding kinetics. However, the kinetics of Skp-uOMP binding is on the order of

milliseconds at ~10^2 nM, and fast enough to prevent uOmpC aggregation at ~1 μM (Wu, Xi et al., 2011). This result shows that Skp certainly possesses AA for uOmpC. This fast binding led Lyu et al. to conclude that uOMPC diffuses rapidly into the Skp trimer (Lyu, Shao et al., 2012). However, this picture of Skp as a static structure raises a second question: *Is Skp a constitutive trimer?*

Although Skp showed AA with uOMPC, holding chaperone activity (section 1.2) requires that Skp also accelerate folding. The best evidence for folding acceleration comes from two studies, by Entzminger et al. and McMorran et al. (Entzminger, Chang et al., 2012, McMorran, Bartlett et al., 2013). Entzminger et al. showed that Skp possesses AA for lysozyme and two antibody fragments. For those three clients, AA correlated with accelerated refolding. This correlation led Entzminger et al. to conclude that Skp is a "holdase" (i.e., a holding chaperone, see section 1.2). In the second study, McMorran et al. (2013), used fluorescence to study Skp CA with the uOMP PagP. PagP fluorescence kinetics showed acceleration of PagP folding in the presence of Skp. However, this acceleration depended on the presence of PG headgroups in the membranes (McMorran, Bartlett et al., 2013). This increased folding and PG-requirement led McMorran et al. to propose that Skp CA for uOMPs depends on electrostatics (McMorran, Bartlett et al., 2013).

Other results call the CA of Skp into question. Thoma et al. (2015) used atomic force microscopy (AFM) to assay the CA of SurA and Skp. Whereas SurA aided AFM-refolding of the uOMP FhuA, Skp had no effect. However, Thoma and Burmann et al. measured slow dissociation of uOMPs from Skp, on the time scale of hours (Burmann, Wang et al., 2013, Thoma, Burmann et al., 2015). Wheras these slow off rates support

AA, holding CA also requires fast release to the membrane (see section 1.2 and figure 1.5, k_{ci}). Additional studies show mixed results for Skp CA. Measurement of OmpA folding from Skp shows weak CA that depends on lipopolysaccharide (LPS) or BamA (Bulieris, Behrens et al., 2003, Patel and Kleinschmidt, 2013). Hoever, without LPS or BamA, Skp inhibits OmpA folding (Bulieris, Behrens et al., 2003, Patel and Kleinschmidt, 2013).

To summarize, Skp rapidly and tightly binds disordered protein (Qu, Mayer et al., 2007, Wu, Ge et al., 2011, Lyu, Shao et al., 2012, Moon, Zaccai et al., 2013, De, Jeong et al., 2014), has holding activity for some soluble proteins (Entzminger, Chang et al., 2012), and has AA for uOMPC and lysozyme (Entzminger, Chang et al., 2012, Lyu, Shao et al., 2012). However, without LPS or BamA, Skp inhibits OmpA folding (Bulieris, Behrens et al., 2003, Patel and Kleinschmidt, 2013), and shows no CA with FhuA (Thoma, Burmann et al., 2015). Thus, there is a question about Skp CA: *Is Skp specialized as an antiaggregation (AA) QC protein, or is it a holding chaperone for uOMPs*?

1.4 Common Patterns in the Proteostasis Networks of the Cytoplasm and Periplasm

If Skp has a specialty, then it cooperates with a network of proteins. A proteostasis network is a set of molecules with QC activities (e.g., degradation, AA, foldase activity, see above) that coexist and interact in the same compartment. Thus, for *E. coli*, two networks exist, one in the cytoplasm and one in the periplasm. Comparing these networks will highlight a functional need that Skp could uniquely fulfill.

First, both networks include large (~1 megadalton) cage-forming complexes with holding activity or degradation activity. The periplasm has one such complex, DegP,

which is a holding chaperone and protease that degrades uOMPs (Hansen and Hilgenfeld, 2013). The cytoplasm has three such complexes, chaperonins I and II (GroEL/GroES and CCT) and the proteasome (Young, Agashe et al., 2004, Xie, 2010, Millan-Zambrano and Chavez, 2014).

Second, both networks include proteins that have holding activity and foldase activity (see above, and figure 1.5). In the cytoplasm, these are chaperonins, trigger factor and HSP70 (Walter and Buchner, 2002, Young, Agashe et al., 2004). In the periplasm, this holding/foldase function is fulfilled by SurA, FkpA, and Spy (Helbig, Patzer et al., 2011, Thoma, Burmann et al., 2015, Stull, Koldewey et al., 2016).

Finally, both networks have proteins with specialized AA or holding activity. In the cytoplasm of bacteria, sHSPs have AA and adapt structurally to clients (Haslbeck and Vierling, 2015). In the cytoplasm of eukaryotes, the prefoldin complex is a holding chaperone with a structure that is similar to Skp (Walton and Sousa, 2004). In the periplasm, Skp is currently believed to be a dedicated holding chaperone (see above).

Despite these similarities, there are two key differences between the cytoplasmic and periplasmic networks. Most importantly, the periplasm does not contain ATP as an energy source. Without an energy source, disaggregase activity seems highly improbable (Aguado, Fernandez-Higuero et al., 2015). Thus, aggregation is likely to be irreversible in the periplasm, and poses a special hazard for *E. coli*. Moreover, the OM is open to ions and other small molecules, which may interfere with uOMP folding.

This comparison highlights a need for a dedicated AA factor in the periplasm. Skp may play the same role as cytoplasmic sHSPs or prefoldin (Stirling, Bakhoum et al., 2006,

Haslbeck and Vierling, 2015). If so, Skp should possess the sHSP traits of structural adaptability and AA (Haslbeck and Vierling, 2015).

This thesis shows that Skp possesses these traits, by addressing the three questions posed in section 1.3. The first question, do free uOMPs contract to fit inside this cavity, or does Skp adapt to them (figure 1.7), is addressed in chapter 2. The second question, is Skp a constitutive trimer, is addressed in chapter 4. Finally, the question of whether Skp is specialized as an antiaggregation quality control protein for uOMPs is examined in chapters 3 and 5.

1.5 Overview of Thesis

Chapter 2

Deuterium Labeling Together with Contrast Variation Small-Angle Neutron Scattering Suggests How Skp Captures And Releases Unfolded Outer Membrane Proteins

To understand how Skp binds and adapts to uOMPs of differing structure, we utilized contrast variation small-angle neutron scattering (Zaccai, Sandlin et al., 2015) to measure the radius of gyration (R_g) of Skp alone and in complex with uOmpA and uOmpW. We found that the helical domains of Skp are expanded in solution relative to the apo crystal structure. Contrast variation also enabled us to estimate the relative locations of the center of masses (CMs) of uOMPs with respect to Skp in the complexes. SANS data reveal varying degrees of movement and collapse of the Skp α-helical domains in the complexes relative to apo-Skp. The position of the unfolded barrel domains of uOmpW and uOmpA also vary in the complex, and the folded uOmpA periplasmic domain extends from Skp in multiple orientations. The contrast variation SANS approach provided structural information that cannot be acquired by other methods.

Chapter 3

Skp Chaperones uOmpW, but not uOmpLa or uBamA in Conditions That Promote Aggregation

To determine whether Skp is a holding chaperone for uOMPs, I used heat modifiability and SDS-PAGE delayed-folding assays to test Skp chaperone activity (CA) with the three OMPs of differing size, OmpW, OmpLa, and BamA. With 80% diC_{10}PC:20% diC_{10}PG LUVs, and conditions of unfolding stress (1 M urea), I find that a 70% active population of Skp shows CA with OmpW, but not OmpLa or BamA. For OmpLa, Skp inhibits folding. For BamA, Skp has no effect. CA thus diminishes with uOMP size in these conditions. From these results, I conclude that the Skp CA is limited and varies significantly with the uOMP substrate.

Chapter 4

Skp Trimer Formation is Insensitive to Salt Concentration in the Physiological Range

Using sedimentation equilibrium, we discovered that Skp is not an obligate trimer under physiological conditions and that Na^{+}, Cl^{-}, Mg^{2+}, and $HPO_4^{1-/2-}$ ions are not linked to Skp trimerization. Our results demonstrate that Skp monomers are populated at biologically relevant concentrations, which raises the idea that kinetic formation of Skp-uOMP complexes likely involves Skp monomer assembly around its substrate. In addition, van't Hoff analysis of Skp self-association does not support a previously proposed coupled folding and trimerization of Skp.

Chapter 5

Skp is More Like a Buffer Than a Chaperone for Outer Membrane Proteins

In chapter 4, I proposed a "switching hypothesis", that Skp trimers provide antiaggregation (AA) or holding chaperone activity (CA) whereas Skp monomers do not. To test this idea, I designed two constructs, FAI and SAI, with reduced ability to form trimers. SE data showed FAI does not form trimers, whereas SAI trimers are weakened by 1.8 kcal mol^{-1}. With a new SV AA assay I test the switching hypothesis by comparing the AA of FAI to wild-type Skp. I find that monomeric FAI has much weaker AA than wild-type Skp. Also, preliminary fluorescence binding data suggest FAI loses activity due to loss of affinity for uOmpLa. Taken together, these results support the switching hypothesis, that Skp function is activated by increased expression.

To determine whether Skp AA varies with substrate size, I measured AA for uOmpA, uOmpLa, and uBamA-TM. AA of Skp with uOMPs did not correlate with uOMP size, R_H, or any other parameter calculated from primary structure. With this AA data for three uOMPs, I asked whether Skp is a holding chaperone for uOMPs. Using SDS-PAGE delayed folding assays with diC11PC LUVs, I found that Skp AA does not correlate with chaperone activity (CA). This result suggests that Skp is not a holding chaperone for uOMPs, but a specialized AA factor.

Finally, I asked whether the affinity of Skp for uOMPs matches the concentration at which uOMPs aggregate. To make this measurement, I developed an assay to measure uOMP aggregation using intrinsic tryptophan fluorescence. I found that uOmpLa and uBamA aggregate at ~10 nM, the concentration at which Skp binds uOMPs. Based on

these results I suggest a shift in paradigm: that AA is the primary function of Skp, with CA a secondary, minor function.

1.6 Figures

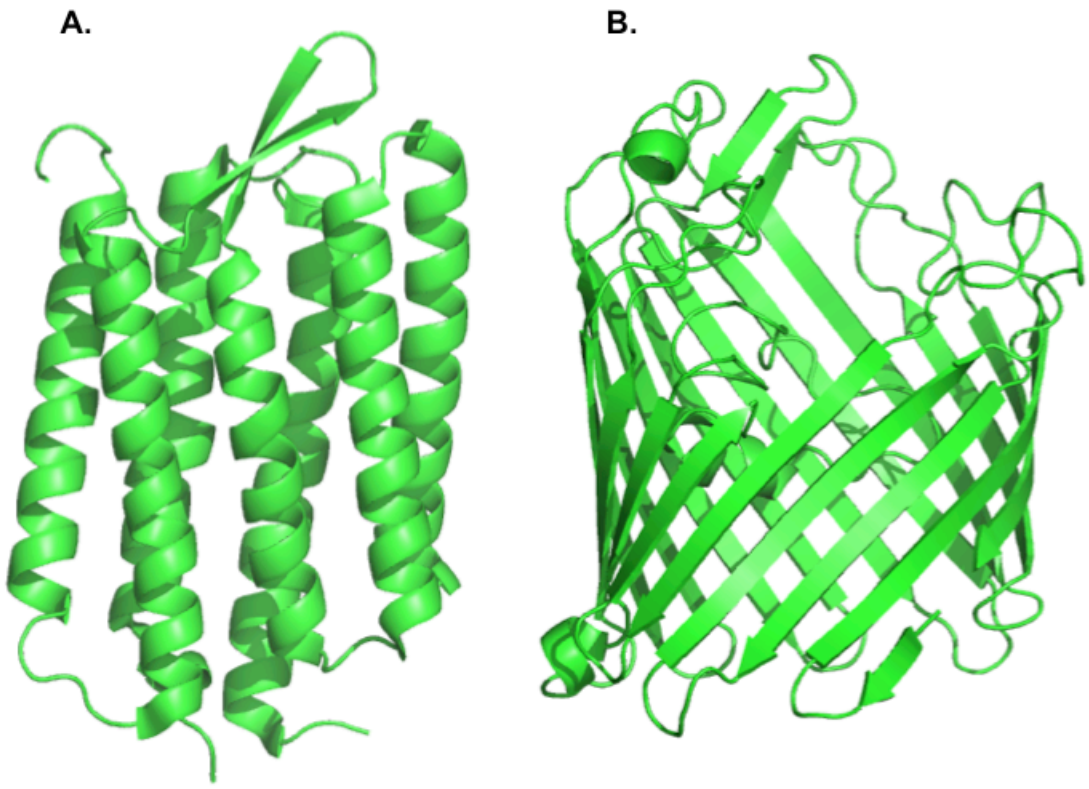


Figure 1.1. Characteristic α-helical and β-barrel Membrane Proteins. A., A ribbon Crystal structure (PDB ID 4XXJ) of Bacteriorhodopsin from *Halobium salinarium* (Bratanov, Balandin et al., 2015). α-helical proteins make up much of the plasma membrane of bacteria. B., Crystal structure (PDB ID 2OMF) of outer membrane protein F from *E. coli* (Cowan, Schirmer et al., 1992). β-barrels make up most of the mass and volume of the outer membrane of Gram-negative bacteria (Jaroslawski, Duquesne et al, 2009).

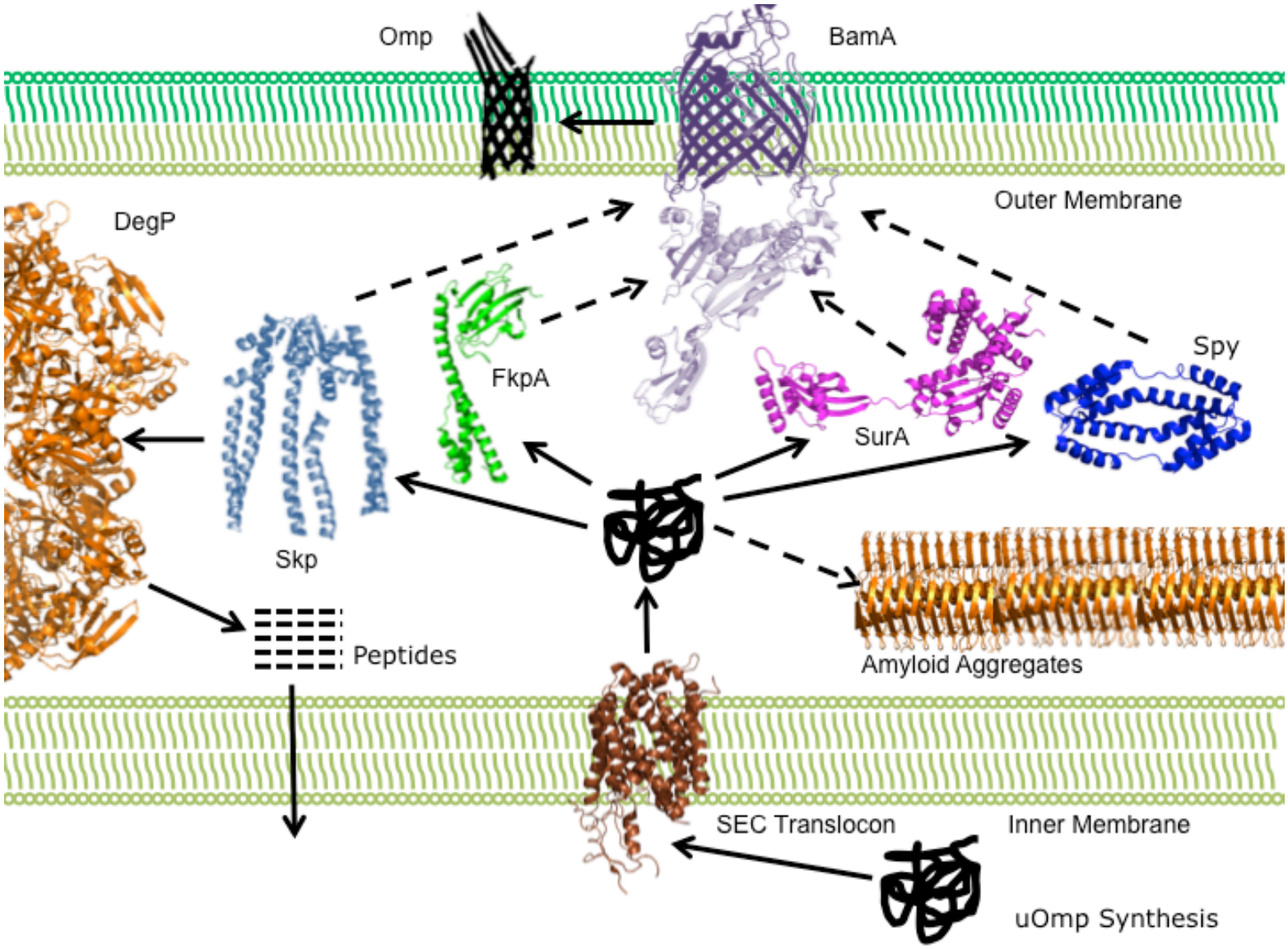

Omp
BamA
Outer Membrane
DegP
FkpA
Spy
SurA
Skp
Peptides
Amyloid Aggregates
SEC Translocon
Inner Membrane
uOmp Synthesis

Figure 1.2. A Simplified View of Proteostasis in the *E. coli* Periplasm. uOMPs are synthesized in the periplasm (bottom, dark squiggle), and transported cotranslationally through the SEC translocon. The main four factors that aid uOMP transport are SurA (magenta), Skp (slate), FkpA (green), and Spy (blue). BamA (purple) is essential for most uOMPs to fold into the OM, and works with four accessory proteins BamB, C, D, and E (omitted for clarity). SurA is essential, whereas FkpA responds to heat, and Spy to tannins. DegP degrades uOMPs. Amyloid aggregates may occur under stress conditions, or when Skp is genetically deleted. Solid arrows, interactions that have been shown by experiment. Dotted arrows, proposed interactions for which there is less evidence.

The PDB IDs of the structures used to create this figure are: SEC Translocon, 2ZQP, DegP, 3OTP, Omp, Ompx structure 1QJ8, FkpA, 1Q6U, SurA, 1M5Y, Spy, 3O39, Skp, a model from (Zaccai, Sandlin et al., 2015), BamA, 5D0O, Amyloid Aggregates, 2N0A.

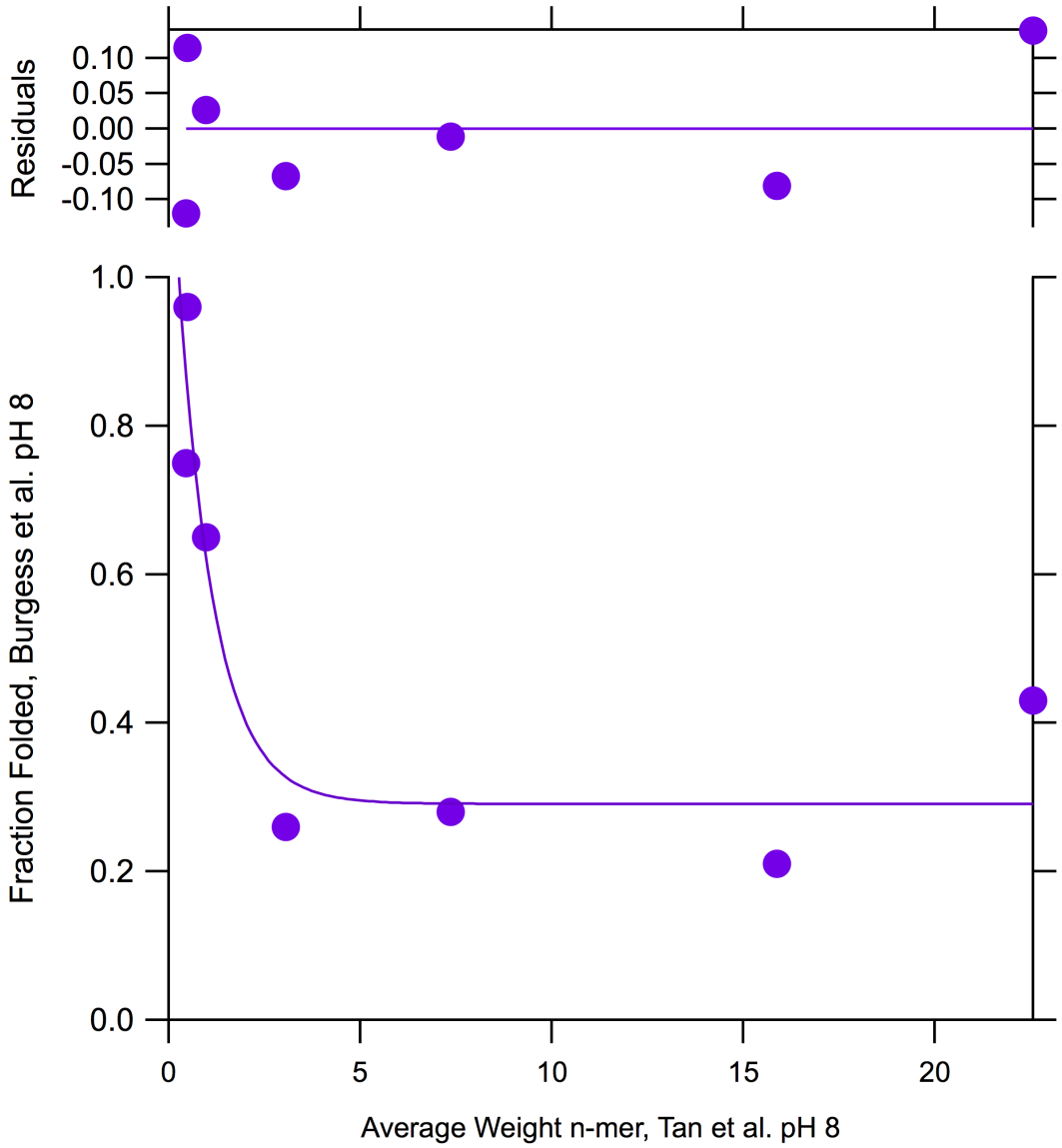


Figure 1.3. Folding Yield Drops Exponentially with Aggregate Yield in the Absence of Lipids. Folding efficiency of seven different uOMPs at pH 8.0, in the presence of 1 M urea and diC_{12}PC LUVs, plotted as a function of the molecular weight of uOMP aggregates without LUVs (Burgess, Dao et al., 2008, Ebie Tan, Burgess et al., 2010). FadL was not included in this analysis, because folding was never detected in the Burgess et al. pH 8.0 conditions.

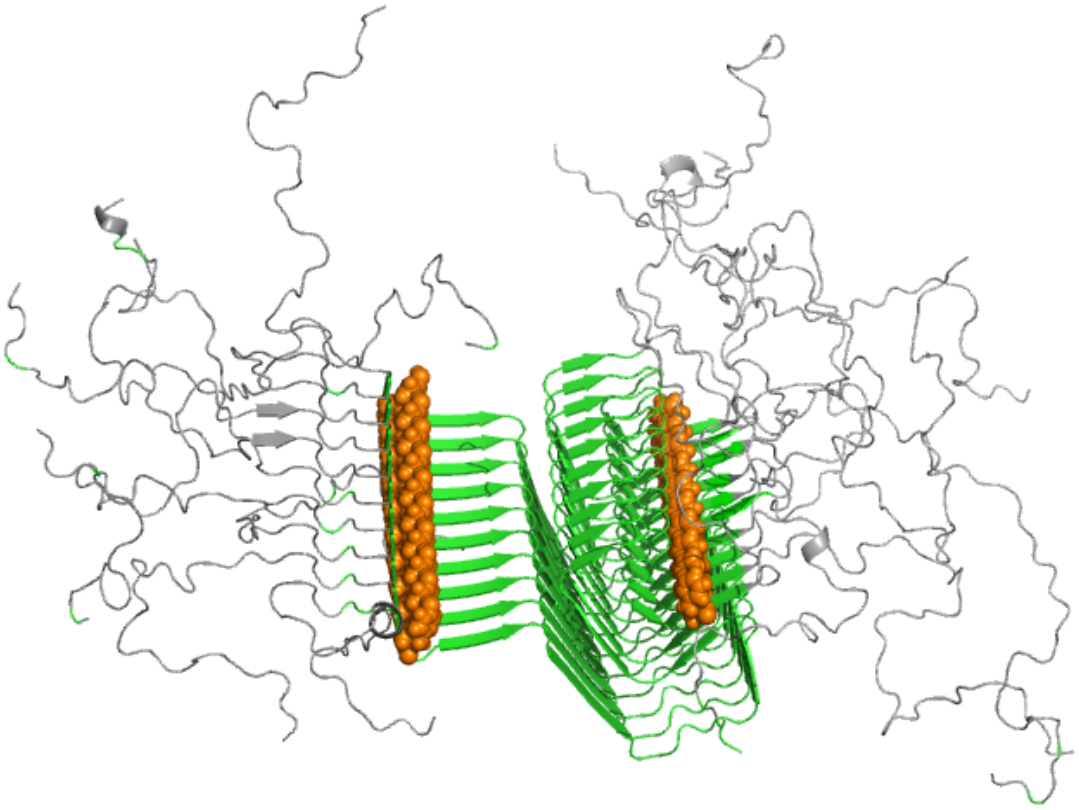


Figure 1.4. The Structure of Amyloid Fibrils Formed by α-Synuclein. Green, the repeating arrangement of highly ordered parallel beta-sheet structure is the key feature of amyloid. Orange spheres, aligned stacks of aromatic side-chains. PDB ID 2N0A (Tuttle, Comellas et al., 2016).

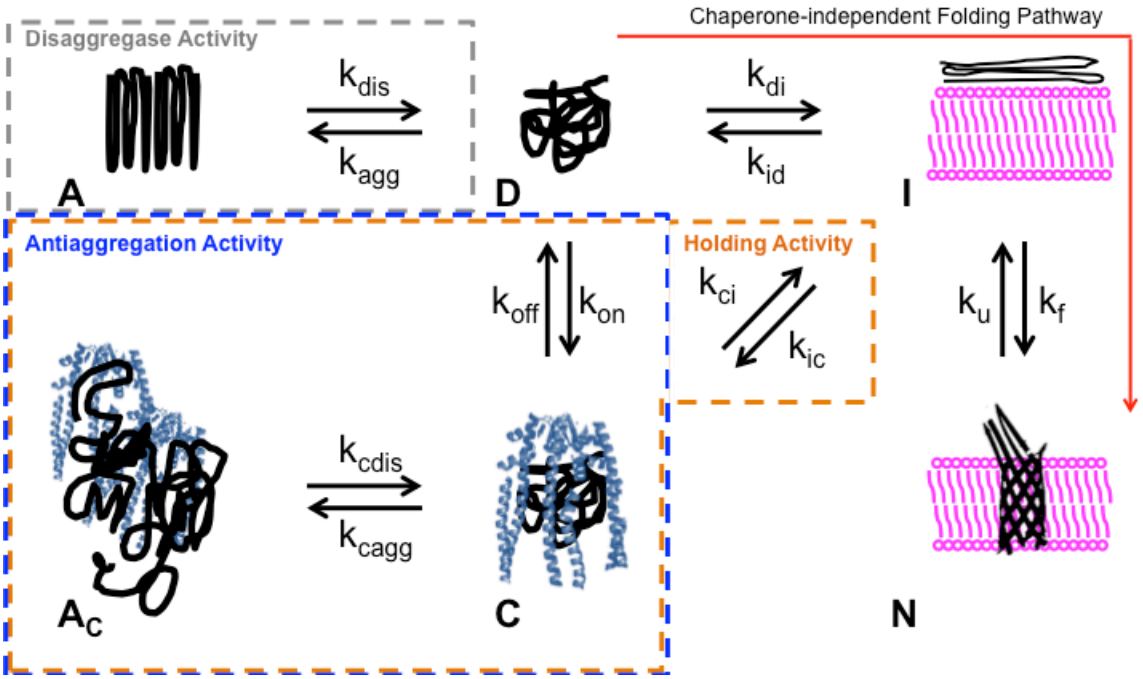


Figure 1.5. A Simplified Kinetic Description of Antiaggregation Activity and Chaperone Activity for OMP Folding. Grey dashed rectangle, disaggregase activity increases the rate k_{dis} from aggregated, A, to denatured, D states. Blue dashed rectangle, antiaggregation activity (AA) requires that binding, k_{on}, of D to a protein to form a complex C is faster than k_{agg}, the rate of aggregation. Also, k_{cagg}, the rate of production of bound aggregates, A_c, must be slower than their rate of dissociation, k_{cdis}. Orange dashed rectangle, holding activity requires AA and fast uOMP release, k_{ci}, to the membrane interface, I. For holding activity to work, the path through the chaperone must be faster than the pathway without it (red arrow). Foldase activity (not shown), works by accelerating the folding rate, k_f, while bound to the chaperone.

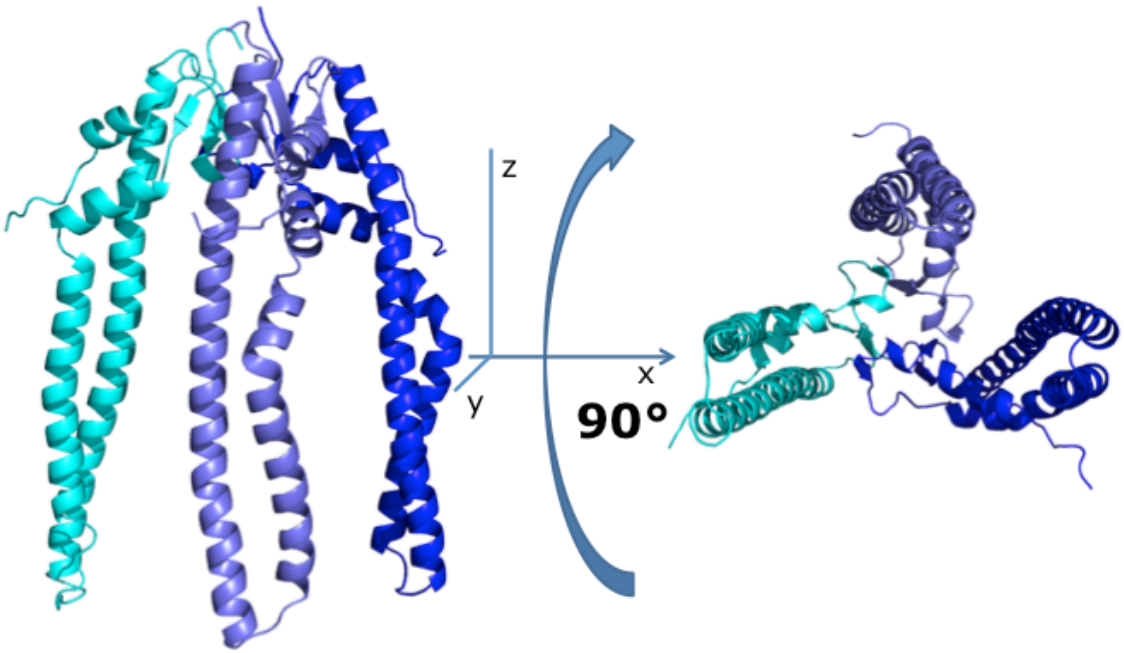


Figure 1.6. A Ribbon Representation of the Skp Structure in Solution. This model is representative of the apo-Skp ensemble simulated in Zaccai, Sandlin et al. (2015).

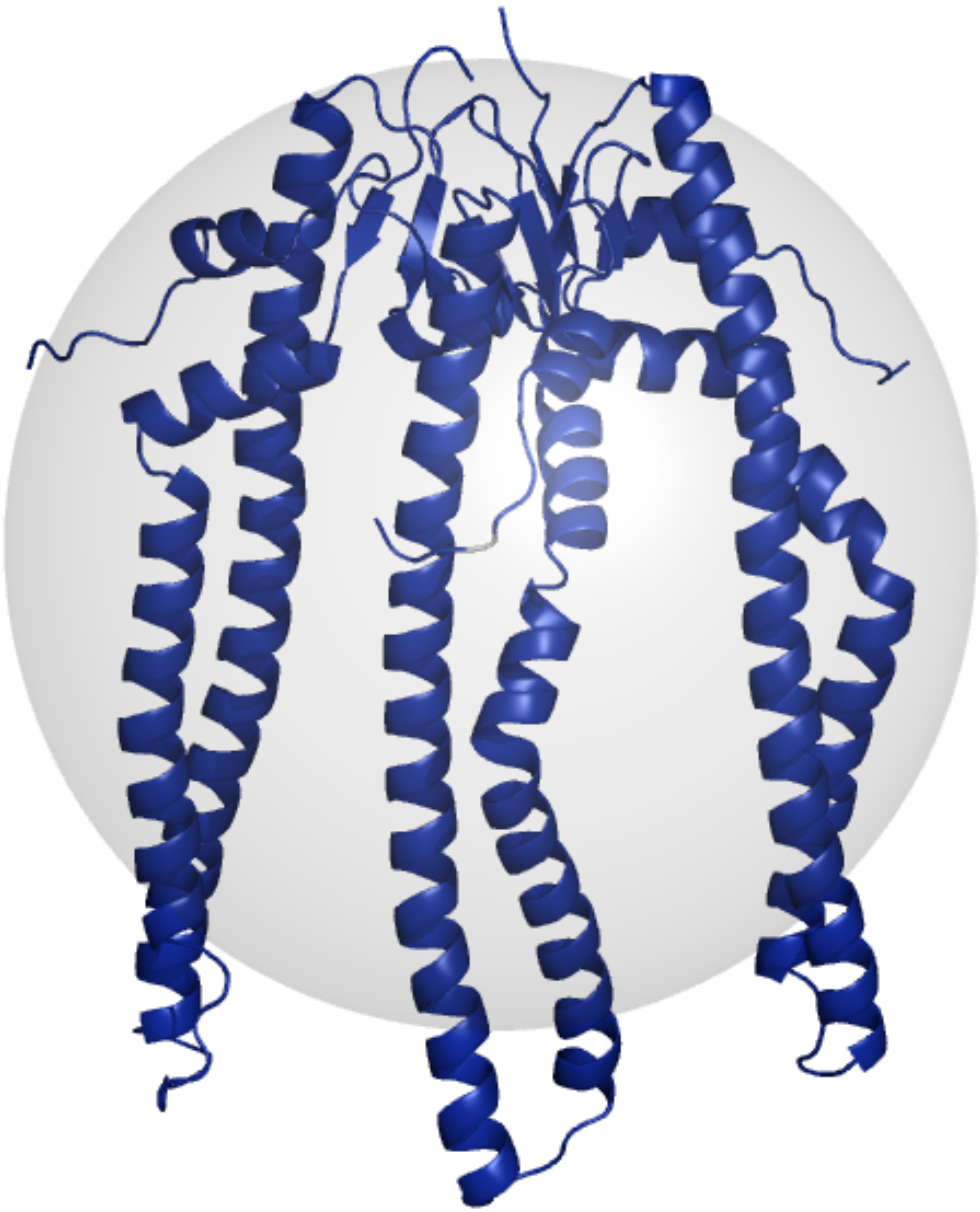


Figure 1.7. The Paradox of the Skp Cavity. Skp-uOmpX complexes form spontaneously by rapid dilution from 8 M urea *in vitro*. Light grey, the estimated hydrodynamic radius of 38 Å for uOmpX (Wilkins, Grimshaw et al., 1999). Blue, the structure of Skp forms a small 21 Å cavity (Burmann, Wang et al., 2013). This discrepancy suggests an obvious question: How do OMPs get inside, and, once inside, how do they get out?

Chapter 2

Deuterium Labeling Together with Contrast Variation Small-Angle Neutron Scattering Suggests How Skp Captures and Releases Unfolded Outer Membrane Proteins

Nathan R. Zaccai*[,1,2], Clifford W. Sandlin*, James T. Hoopes[†], Joseph E. Curtis[‡], Patrick J. Fleming*, Karen G. Fleming*, Susan Krueger[‡,1]

*T.C. Jenkins Department of Biophysics, Johns Hopkins University, Baltimore, Maryland, USA

[†]Institute for Bioscience and Biotechnology Research, University of Maryland, Rockville, Maryland, USA

[‡]National Institute of Standards and Technology, Gaithersburg, Maryland, USA

[1]Corresponding authors: e-mail address: nrz20@cam.ac.uk; susan.krueger@nist.gov

[2]Current address: Cambridge Institute for Medical Research, University of Cambridge, Hills Road, Cambridge CB2 0XY, United Kingdom.

Originally published in Methods in Enzymology, Volume 566, pages 159-210, January 23rd, 2016

Preface: Statement About My Role as the Second Author of this Publication

I worked closely with Dr. Nathan Zaccai on the expression and purification of the OmpW

and Skp protein, worked directly with him on assembly of the Skp-OmpW complexes, and assisted him on-site at NIST with size-exclusion chromatography of the Skp-OmpW complex. In addition, I aided in data collection on the beamline and performed the Guinier and Stuhrmann analyses in parallel with Dr. Nathan Zaccai, Dr. Karen Fleming, and Dr. Susan Kreuger to ensure correctness of fitting and calculations of R_g and D_{CM}. I also provided advice to both Dr. Zaccai and Dr. Patrick Fleming regarding consideration of potential issues with uniqueness in least-squares fitting using molecular models. In addition to writing the sections on expression and purification, I provided editing advice and critique of the manuscript. I agree with the statements in the paper, and include it in my thesis dissertation, almost completely in the form it was published in Methods in Enzymology as relevant to the question of understanding how Skp adapts to uOMP substrates that vary in structure (Zaccai, Sandlin et al.). Small alterations have been made in formatting. All instances of "Fig(s)." or "Eq(s)." have been altered to read "Figure(s)" or "Equation(s)". Also all instances of "Table" in sentences or parentheses were altered to "table", and all instances of "beta-barrel" were changed to "β-barrel". Subsections, citations, and headings were organized to follow the overall format of the thesis chapters. These formatting changes do not change the presentation, content, or findings of the study.

2.1 Introduction

In Gram-negative bacteria, unfolded outer membrane β-barrel proteins (uOMPs) interact with periplasmic chaperones in order to be trafficked to the outer membrane. Two of these chaperones, Skp and SurA, form stable complexes with uOMPs as they travel

across the aqueous periplasm. These chaperones exhibit holdase activity because they protect their substrates from aggregation, as well as from cleavage by the periplasmic protease DegP (Wu, Ge et al., 2011).

Even though Skp and SurA appear to have redundant functions, their relative activities are dependent on bacterial species and whether or not the bacteria are under stress (Rhodius, Suh et al., 2006 West, & Gross, 2006, Sklar, Wu et al., 2007 & Silhavy, 2007, Denoncin, Schwalm et al., 2012Silhavy, & Collet, 2012). Although not a lethal mutation in *Escherichia coli*, loss of Skp results in a relative decline in uOMP concentration in the bacterial outer membrane (Chen and Henning, 1996).

Skp binds a wide range of bacterial uOMPs with nanomolar affinity (Qu, Mayer et al., 2007, Holst, & Kleinschmidt, 2007, Moon, Zaccai et al., 2013, Gessmann, & Fleming, 2013). At least 19 different *E. coli* uOMPs interact with Skp (Jarchow, Luck et al., 2008 & Skerra, 2008). These substrates vary in sequence composition and in size (20–150 kDa). In bacteria, Skp–uOMP complexes form in the periplasm near the bacterial inner membrane (Schäfer, Beck et al., 1999). The Skp–uOMP complex is then transported to the outer membrane where the uOMP folding is thought to be mediated by the Bam complex (Webb, Heinz et al., 2012). *In vivo*, the Skp-presented bound uOMP could directly interact with the Bam complex, or uOMP could first be delivered by Skp to SurA, which then transports it to the Bam complex (Sklar, Wu et al., 2007, Ieva, Tian et al., 2011 & Bernstein, 2011, Schwalm, Mahoney et al., 2013 & Silhavy, 2013). Skp may also transfer the uOMP directly to the bacterial outer membrane as *in vitro* experiments demonstrate that Skp-bound uOMPs retain the ability to fold into lipid bilayers containing phosphatidylcholine and phosphatidylglycerol (Bulieris, Behrens et al., 2003,

McMorran, Bartlett et al., 2013).

The holdase activity of Skp is not limited to specific membrane proteins (Jarchow, Luck et al., 2008). Skp is able to form a complex with the periplasmic (PP) domain of the auto transporter EspP (Ieva, Tian et al., 2011). It can also inhibit the aggregation and assist the folding of a number of soluble proteins.

Notable examples include single-chain antibodies (Entzminger, Chang et al., 2012) and lysozyme (Walton and Sousa, 2004). Skp forms a stable trimer (50 kDa) in solution as determined by gel filtration as well as cross-linking methods (Schlapschy, Dommel et al., 2004) and as supported by crystal structures (Korndörfer, Dommel et al., 2004, Walton and Sousa, 2004). Three α-helical "tentacles" extend out from the β-strand Skp "body" to create a cavity sufficiently large to surround a 25-kDa substrate. This domain architecture, termed jellyfish-like fold, had been previously described for other holdases, including the archaeal and eukaryotic prefoldins and the mitochondrial small Tim proteins (Stirling, Bakhoum et al., 2006). Like Skp, these holdases have the common property of protecting their substrates from aggregation. However, no sequence identity to Skp was present, and unlike the trimeric Skp, these proteins are heterohexamers, with six α-helical tentacles each.

X-ray and subsequent NMR data could not directly provide an explanation for the ability of Skp to bind a wide range of uOMPs of different sizes. It was hypothesized that an uOMP transmembrane (TM) domain is bound within an adaptable cavity formed by the α-helical tentacles of Skp. NMR analysis of Skp–OmpA and Skp–OmpX complexes indicated that the uOMP TM region contains little secondary structure while in complex with Skp (Walton, Sandoval et al., 2009, Burmann and Hiller, 2012, Burmann, Wang et

al., 2013, Callon, Burmann et al., 2014). However, the ability of Skp to bind diverse substrates begs the question as to whether changes occur in the tertiary structure of Skp upon binding a client substrate.

In this study, a series of contrast variation small-angle neutron scattering (Rassam, Copeland et al., 2015) experiments were performed on Skp alone in solution and on two different Skp–uOMP complexes: Skp–OmpA and Skp–OmpW, in order to clarify the structural basis of uOMP presentation by Skp. OmpA has a role in *E. coli* cell morphology and stability. Folded OmpA (35 kDa) has a C-terminal PP domain that folds independently from its TM domain (Walton, Sandoval et al., 2009, Danoff and Fleming, 2011). OmpW (21 kDa) is an integral membrane protein that is required for resistance to phagocytosis (Wu, Ge et al., 2011). The atomic resolution structures of the TM domains of both these proteins revealed an eight-stranded β-barrel when folded into membranes (Pautsch and Schulz, 1998, Hong, Patel et al., 2006 & van den Berg, 2006). The folding properties of these proteins have been extensively investigated (Burgess, Dao et al., 2008, Moon, Zaccai et al., 2013). Selective labeling by deuteration of OmpW and OmpA in the Skp–uOMP complexes allowed the determination of the individual structures of uOMP and Skp as well as their relationship to each other in complexes.

The SANS data indicate that Skp interacts with the uOMP TM domain in a manner analogous to the binding mechanism of prefoldins binding their unfolded substrates (Martin-Benito, Boskovic et al., 2002, Martin-Benito, Gomez-Reino et al., 2007). The bulk of the unfolded TM domain of both OmpW and OmpA is within the Skp cavity but some parts must be on the exterior of Skp near the tips of the tentacles to be in agreement with the SANS data. The PP domain of OmpA can assume a number of

different positions outside the Skp cavity and still be consistent with the data. These results yield the first direct structural evidence that the α-helical Skp tentacles move closer together on binding its substrate and that the structure of the individual helices of the tentacles is different when binding different uOMPs. Importantly, the SANS data represent unique information that could not be obtained without deuterium labeling of the uOMPs. Analysis of the SANS data provides experimental support for a simple clamp-like mechanism used by jellyfish-like chaperones (Stirling, Bakhoum et al., 2006). It also allows postulation of a mechanism of Skp binding and delivery of uOMPs to the bacterial outer membrane.

SANS From Biological Molecules In Solution

SANS is able to provide the size, molecular mass, and shape of a macromolecular complex in solution on length scales between approximately 10 Å to about 1000 Å (Beaudry, Petersen et al., 1976, Moore, 1982, Jacques and Trewhella, 2010, Zaccai, 2012). While analytical ultracentrifugation (AUC) can be used to ascertain the Skp–uOMP stoichiometry and small-angle X-ray scattering (SAXS) can yield the overall size and shape of the Skp–uOMP complex, analysis of contrast variation SANS data can uniquely retrieve the internal structure and organization of the complex. Recent reviews (Heller et al. 2010,Whitten & Trewhella, 2009) as well as classic papers (Ibel and Stuhrmann, 1975, Beaudry, Petersen et al., 1976) describe the contrast variation technique in detail.

Contrast and Scattering Intensity

Because neutrons interact with atomic nuclei, the strength of the scattering interaction is not dependent on the atomic number, Z, of the atom, as is the case for X-rays. Therefore, the light elements such as H, C, N, and O can scatter neutrons just as strongly as heavier elements. Also, neutrons are sensitive to different isotopes of an element, such as hydrogen and deuterium. Since SANS does not provide information on atomic length scales, the strength of the scattering interaction can be described in terms of a uniform scattering length density of the entire molecule, ρ, within the molecular volume, V (in cm^3 or $Å^3$). For biological molecules in solution, the strength of the scattering is further defined as the difference in the scattering length densities of the molecule and the solvent within the same molecular volume. This difference in scattering length densities is known as the contrast, and is written as

$$\Delta\rho = \rho - \rho_s \qquad \text{Equation 2.1}$$

where ρ_s is the scattering length density of the solvent. When the scattering length densities of the molecule and solvent are the same, $\Delta\rho = 0$. This is called the contrast match point. The scattering length densities and contrast are usually expressed in units of cm^{-2} or cm $Å^{-3}$, but can be found stated in units of $Å^{-2}$.

The measured SANS intensity from a solution of monodisperse, randomly oriented biological macromolecules can be written in terms of the contrast as

$$I(q) = n(\Delta\rho)^2 V^2 \langle |F(\vec{q})|^2 \rangle \qquad \text{Equation 2.2}$$

where $|F(\vec{q})|^2$ is the scattering form factor, which depends on the shape of the molecule and n is the number density of molecules (in cm^3). The brackets represent an averaging over all orientations of the molecule. The magnitude of the scattering vector q can be

written as

$$q = \frac{4\pi \sin(\theta)}{\lambda}$$ Equation 2.3

where 2θ is the scattering angle (in degrees), measured from the axis of the incoming neutron beam, and λ is the neutron wavelength. The wavelength is usually expressed in nm or Å, such that q is stated in units of nm^{-1} or $Å^{-1}$. It can be seen from Eq. (2.2) that the scattering intensity at the contrast match point, $\Delta\rho = 0$, is zero.

Radius of Gyration and Forward Scattering Intensity

The radius of gyration, R_g, and the forward scattering intensity, I(0), which is the scattering intensity at q = 0, are two important model-independent parameters that are obtained from SANS data. R_g provides information about the size of the molecule. A shape must be assumed for the molecule to relate R_g to the molecular dimensions. For example, a solid sphere has a radius, r, which is equal to $1.3R_g$.

By definition, the scattering form factor in Eq. (2.2) is equal to 1.0 at q = 0. Thus, I(0) is related to the number density, n, as

$$I(0) = n(\Delta)^2 V^2$$ Equation 2.4

The number density can be written in terms of the concentration of the molecule, c (g cm^{-3}), as

$$n = \frac{cN_A}{M_W}$$ Equation 2.5

where M_w is the molecular weight of the molecule (in Da, where1 Da = 1 g mol^{-1}) and N_A is Avogadro's number. Equations (2.4) and (2.5) can be used to relate I(0) to the M_w of the molecule, if the SANS data are on an absolute scale, usually in units of cm^{-1}

The Guinier approximation (Guiner and Fournet, 1955),

$$I(q) = I(0)\exp\left(-q^2\frac{R_g^2}{3}\right)$$ Equation 2.6

can be used on the low-q portions of the data to obtain values for R_g and I(0). This "low-q" analysis is valid only in the region where $qR_g \leq 1.3R_g$ and I(0) are found by plotting the natural log of Eq. (2.6) such that

$$\ln\big(I(q)\big) = \ln\big(I(0)\big) - q^2\frac{R_g^2}{3}$$ Equation 2.7

A linear fit of ln(I(q)) versus q^2 (Equation 2.7) to the low-q portion of the data allows the determination of R_g from the slope and I(0) from the intercept. If there are aggregates of the molecule in the solution, Eq. (2.7) will not be linear. Rather it will have some curvature and the fit to a straight line will be poor. The effects can be subtle or very obvious depending on the severity of the aggregation (Jacques and Trewhella, 2010). If aggregation is present, whether subtle or severe, the R_g value found from Eq. (2.7) no longer represents that of a single monomer in a monodisperse solution. Rather, it is influenced by the larger aggregates present in the solution.

Another method to obtain R_g, which makes use of all of the data rather than a limited data set at small q values, is to use the distance distribution function, P(r) versus r, to determine R_g and I(0) (Glatter and Kratky, 1982). This function represents the probability distribution of distances, r, between all pairs of atoms in the molecule. The result is a smooth histogram-like plot that peaks at the most probable distance in the molecule. Thus, the shape of the P(r) versus r curve depends strongly on the shape of the molecule.

P(r) is typically obtained from the SANS data using an indirect Fourier transformation method (Glatter, 1977, Moore, 1982, Semenyuk and Svergun, 1991) using the relation

$$I(q) = 4\pi V \int_0^{D_{max}} P(r) \frac{\sin(qr)}{qr} dr \qquad \text{Equation 2.8}$$

This analysis requires a stipulation by the user of a maximum dimension, D_{max}, beyond which P(r)=0. Typically, several values of D_{max} are explored in order to find the range over which the P(r) function does not change as a function of D_{max}. The P(r) function is also sensitive to aggregation. The more severe the aggregation, the more difficult it is to determine D_{max} (Jacques and Trewhella, 2010).

Contrast Variation

The scattering length density of water is shown as a function of $\%D_2O$ in the solvent in figure 2.1, along with scattering length densities of some typical biological molecules and compounds. The lines are straight for CH_2 and CD_2 because there is no exchange of deuterium for hydrogen as the% D_2O increases in the solvent. However, for proteins and nucleic acids, labile hydrogen atoms, i.e., those bound to nitrogen and oxygen, will exchange with deuterium in the solvent, so their scattering length densities vary with increasing % D_2O. A vertical line representing the contrast, $\Delta\rho$, between protein and water is shown on the graph for 10% D_2O in the solvent. Note that the protein and water lines cross at 40% D_2O, which is the contrast match point for most typical proteins. The match points for DNA, RNA, lipid head groups, and CH_2 can be found in the same manner. Note that the scattering length densities for perdeuterated molecules do not cross the water line. Thus, the match points for these molecules are greater than 100% D_2O and cannot be reached in practice.

In cases where a molecule is actually a complex consisting of components that have different scattering length densities, the contrast is different for each of the

components. If the ratio of H_2O to D_2O in the solvent is varied, the contrast of each component will change as a function of the concentration of D_2O in the solvent. Thus, contrast match points exist for each of the two components as well as the entire complex. By varying the amount of D_2O in the solvent, one component can be essentially transparent at its match point while the other is still visible. It is this feature of SANS that makes the method so powerful for selective measurement of individual components within a complex. From figure 2.1, it is clear that proteins and nucleic acids have different contrast match points. The protein match point is around 40% D_2O, meaning that only the DNA or RNA is visible at this contrast. The DNA and RNA match points are around 65% D_2O such that only the protein is visible under these conditions. Therefore, complexes consisting of proteins and nucleic acids are ideal candidates for contrast variation experiments.

For a complex consisting of two proteins, replacement of the nonexchangeable hydrogen atoms, i.e., those bound to carbon, with deuterium in one of the components is required in order for the two match points to be different. This is typically accomplished by expressing one of the protein components using bacterium grown in deuterium-enriched media. Figure 2.1 shows that the match point of perdeuterated proteins, in which all nonexchangeable hydrogen atoms have been replaced by deuterium, is above 100% D_2O. Therefore, partially deuterated proteins are generally used for contrast variation experiments so that the match point of the deuterated component is somewhere between 60% D_2O and 100% D_2O. The exact match point of a deuterated component is dependent on the amount of deuteration achieved. The contrast variation experiment can be used to verify this parameter, especially if a reliable determination cannot be made by other

methods such as mass spectrometry.

Contrast Match Point Analysis of the SANS Contrast Variation Data

For a two-component complex in which the components have different contrasts, the contrast match points for the complex are determined from I(0). Expanding on equation 2.4,

$$I(0) = n(\Delta\rho)^2V^2 = n(f_1\Delta\rho_1V_1 + f_2\Delta\rho_2V_2)^2 \qquad \text{Equation 2.9}$$

The number density, n, is defined as in Eq. (5), but it is now in terms of the concentration and M_w of the entire complex. $\Delta\rho$ now refers to the mean contrast of the entire complex and V is the volume of the complex. The expression on the right is now written in terms of the two components, where f_1 and f_2 are the mass fractions of the 1st and 2nd components in the complex, $\Delta\rho_1$ and $\Delta\rho_2$ are the scattering contrasts of the 1st and 2nd components, and V_1 and V_2 are the volumes of the 1st and 2nd components. Thus, equation 2.9 can be combined with equation 2.5 to relate I(0) to the M_w of the complex if the SANS data are on an absolute scale. This is a good quality assurance test on the data. If the complex has the expected M_w at all contrasts, then it is likely that the stoichiometry of the two components is correct at all contrasts. It is important to recognize that care should be taken to obtain the concentrations of the complexes as accurately as possible at each contrast because uncertainty in this value is a major source of error on the calculation of the M_w.

Because I(0) is proportional to $\Delta\rho_2$ (Equations 2.4 and 2.9), which is in turn dependent on the fraction of D_2O in the solvent, f_{D_2O}, and n is proportional to c (Equation 2.5), the contrast match point of the complex can be determined from the x-intercept of a

linear fit to $\sqrt{I(0)/c}$ versus f_{D_2O}. The match point of the complex can also be calculated from the sequences of the components (Whitten, Cai et al., 2008, Sarachan, Curtis et al., 2013). Using both approaches to determine the match point provides another quality assurance test on the data in that the calculated and experimentally determined match points should agree with each other.

Separation of the Radii of Gyration in a Two-Component Complex

The R_g values obtained at each contrast are related by the relationship (Ibel and Stuhrmann, 1975),

$$R_g^2 = R_m^2 + \frac{\alpha}{\Delta\rho} - \frac{\beta}{\Delta\rho^2} \qquad \text{Equation 2.10}$$

where R_m is the R_g value of the equivalent complex with a homogeneous scattering length density, α is the second moment of the density fluctuations, and β is the first moment of the density fluctuations. For two-component systems with different scattering length densities, the term α relates to the distribution of scattering length densities relative to the center of mass (CM) of the complex, and the term β provides the separation of the scattering CM of the two components (Moore, 1982). A Stuhrmann plot (Ibel and Stuhrmann, 1975) of R_g^2 versus $\Delta\rho^{-1}$ (Equation 2.10) is used to determine R_m, α, and β. If the plot is linear, then β = 0 and the CM of the two components are concentric. In this case, the sign of the slope of the line, α, is an indication of whether the component with the higher scattering length density is on the interior or exterior of the complex. In practice, it is not always easy to distinguish between a linear and a parabolic Stuhrmann plot, especially if α is close to zero and if R_g values are not available close to the individual match points of the two components. Similar information can be obtained from

the parallel axis theorem

$$R_g^2 = f_1 R_1^2 + f_2 R_2^2 + f_1 f_2 D_{CM}^2 \quad \text{Equation 2.11}$$

where R_1 and R_2 are the radii of gyration of the components and D_{CM} is the distance between the scattering CM (Engelman and Moore, 1975). Here,

$$f_i = \frac{\Delta\rho_i V_i}{\Delta\rho V} \quad \text{Equation 2.12}$$

where $\Delta\rho_i$ and V_i refer to the individual components and $\Delta\rho$ and V refer to the complex. The parallel axis theorem provides the radii of gyration of the components directly, whereas they are calculated from the definitions of α and β when the Stuhrmann analysis is used (Whitten, Cai et al., 2008).

Separating the Scattering Intensities in a Two-Component Complex

The scattering intensity from a two-component system with different scattering length densities can be approximated by (Whitten, Cai et al., 2008)

$$I(q) = \Delta\rho_1^2 I_1(q) + \Delta\rho_1 \Delta\rho_2 I_{12}(q) + \Delta\rho_2^2 I_2(q) \quad \text{Equation 2.13}$$

where $I_1(q)$ and $I_2(q)$ are the scattering intensities of components 1 and 2, respectively, and $I_{12}(q)$ is the scattering intensity due to the interference between the two components that occurs because they have different scattering length densities. $I_1(q)$ and $I_2(q)$ are related to the shapes of the two components and $I_{12}(q)$ is related to their spatial distribution. For a given set of measured contrast variation intensities, I(q), and known values for the contrasts, $\Delta\rho_1$ and $\Delta\rho_2$, the three unknowns, $I_1(q)$, $I_2(q)$, and $I_{12}(q)$,are found by solving equation 2.13 simultaneously at all contrasts.

This analysis can be a useful tool to model each component separately and then put them in the proper position with respect to each other in the complex. Data must be

obtained at a minimum of three contrasts to solve for the three unknowns in equation 2.13. However, in practice, successful studies have employed at least five contrast points and high-quality data were obtained at the match points of the individual components. This method worked particularly well in a recent study of a kinase, KinA, in complex with an inhibitor, Sda (Whitten, Jacques et al., 2007).

Structure Modeling

Both SAXS and SANS are being used for structural determination of large protein complexes and for proteins containing flexible regions in solution. Many options are available for modeling multimeric protein complexes using a combination of rigid body and atomistic approaches, as described in recent reviews (Putnam, Hammel et al., 2007, Rambo and Tainer, 2010). The SASSIE software suite (Curtis, Raghunandan et al., 2012) is one tool that is available to assist in the atomistic and rigid body modeling of the structures of biological molecules for comparison to SAXS and SANS data. SASSIE allows users of these techniques access to molecular dynamics (MD), Monte Carlo, docking, and rigid body modeling methods to assist in structure modeling and assessment of how well the models fit the data. Constraints can be incorporated from other techniques such as NMR and AUC. For example, SASSIE has been used for the structure modeling of intrinsically disordered monomeric proteins (Curtis, Raghunandan et al., 2012), large protein complexes (Krueger, Shin et al., 2011, Krueger, Shin et al., 2014, Rubinson, & Kelman 2014), and single-stranded nucleic acids (Peng, Curtis et al., 2014). It has also been applied to the study of monoclonal antibodies using free-energy analysis (Clark, Zhang et al., 2013). A Web version is available (SASSIE-web:Beta, n.d.) for ease

of access and to handle the intensive computational requirements of the structural modeling and data analysis. For a two-component complex, SANS and contrast variation experiments provide the added structural information from the individual components as constraints for modeling the entire complex. If obtainable, the scattering intensities of the separate components (Equation 2.13) can be helpful for the modeling of the individual components and for construction ofthe model structure for the entire complex (Whitten, Cai et al., 2008). However, the R_g and CM distance constraints found by the Stuhrmann (Equation 2.10) and parallel axis theorem (Equations 2.11 and 2.12) analyses add unique information that can be used in the modeling process even in the absence of the component scattering intensities. Often, structural information for one or both of the components alone in solution is used as a starting point for their structures in the complex.

Structure models that fit the SANS data at all contrasts take full advantage of the information content of the contrast variation data set. Whether or not models are constructed from the scattering intensities of the separate components, the model structures should always be judged against the entire contrast variation data set. Working model structures are tested against the data by calculating SANS curves from the model structures and comparing them to the measured SANS curves at all contrasts. The model structures are then revised as necessary to obtain the best global fit to the entire SANS contrast variation data set.

2.2 Materials and Methods

uOMP and Skp Expression and Purification

The uOMPs were cloned and expressed to inclusion bodies as previously

described, except that no detergent was included in the inclusion body washes (Burgess, Dao et al., 2008). Deuteration of the uOMPs was accomplished at the NIST/University of Maryland Biomolecular Labeling Laboratory (BL_2). To produce the deuterated uOMPs, the OMP genes were recloned into the kanamycin-resistant pet28 vector. Expression was performed with HMS cells in deuterated M9 media, containing 60% D_2O. After OD600 = 0.6 was reached, expression was induced with 1 mmol L^{-1} (mM) IPTG and the cells were allowed to grow overnight at room temperature. Deuterated inclusion bodies were prepared with the same buffers as for the hydrogenated inclusion bodies. Inclusion bodies were subsequently stored at -20 °C. The inclusion bodies were dissolved in 20 mM Tris (pH 8), 8.0 M urea, and, after centrifugation at 18,000 rpm (1 rpm = 1/60 Hz) in a Beckman J2-MI with a type 21 rotor for 1 h, the clarified supernatant was stored at -80 °C until use. The uOMP concentrations were determined using their calculated extinction coefficient. The respective coefficients used were 52,955 $M^{-1}cm^{-1}$ for OmpA and 39,420 $M^{-1}cm^{-1}$ for OmpW.

The expression and purification of Skp has been previously described (Moon, Zaccai et al., 2013). Briefly, the *E. coli* skp gene (Gene ID 944861) was amplified by using the following primers: AGGAGATATACCATGGCTGACAAAAT and GTGATGGTGATGTTTAACCTGTTTCA, and was inserted by ligation-independent cloning into the pOPINE expression vector (Berrow, Alderton et al., 2007). The resultant plasmid (NZ100) was then transformed into BL21 (DE3) pLysS *E. coli* (Novagen) for expression of Skp with a C-terminal six-histidine tag. Typically, 2XYT media supplemented with 1% glucose was inoculated from a frozen cell stock and grown overnight at 37 °C. After 1/50 dilution into fresh media, Skp expression was induced 3 h

later with 1 mM IPTG and continued at room temperature for 20 h. Cells were harvested by centrifugation and stored at -20 °C for future use. The frozen cell pellet was resuspended in buffer A (50 mM Tris, pH 8, 500 mM NaCl, 10% (v/v) glycerol, and 20 mM imidazole) with one complete EDTA-free protease inhibitor tablet (Roche) per 50 mL buffer and DNaseI (Sigma). Cells were lysed by sonication and cell debris removed by centrifugation at 19,000 rpm in a size 21 rotor in a Beckman J2-MI centrifuge. The supernatant was applied to a pre-equilibrated Nickel Sepharose High-Performance column (GE Healthcare) and washed multiple times with buffer A (without glycerol). Then, recombinant protein was eluted by addition of a mixture of 50% (v/v) buffer A and 50% (v/v) buffer B (50 mM Tris, pH 8, 500 mM NaCl, and 500 mM imidazole). Pooled fractions were subjected to gel filtration on a Superdex 200 10/300 GL column (GE Healthcare) and equilibrated in GF buffer (20 mM Tris, pH 8, and 200 mM NaCl) for SANS measurements. When required, the SANS buffer was also prepared with D_2O instead of H_2O. Because the 1 M Tris (pH 8) stock did not contain deuterium, this buffer contained 98% (v/v) D_2O and 2% (v/v) H_2O (98% D_2O buffer). The Skp trimer concentration was determined using a calculated extinction coefficient (ε_{280} = 4470 $M^{-1}cm^{-1}$). Nominal Skp concentrations employed were 5.3 mg mL^{-1} in H_2O (0% D_2O) buffer and 3.6 mg mL^{-1} in 98% D_2O buffer.

Formation of Skp–uOMP Complex

The concentration of Skp and uOMP were adjusted so that the uOMP in urea was ≈10 times more concentrated than Skp trimer in GF buffer. The Skp–uOMP complex was then assembled by dropwise addition of the uOMP solution until the first signs of

precipitation were observed. The final urea concentration was estimated to be less than 1 M. At each addition of uOMP, the solution was checked for aggregation. After 30 min incubation at room temperature, the complex sample was repeatedly diluted twofold with GF buffer every 10 min, until the final urea concentration was under 0.1 M. Any aggregates were removed with a 0.22-μm filter and the protein was concentrated to between 2 and 5 mg mL^{-1} with an Ultra-15 Ultracel 30k Centrifugal Filter (Millipore UFC903024) prior to being purified on a Superdex 200 10/300 GL column (GE Healthcare) equilibrated in GF buffer. Fractions containing the Skp–uOMP complex were identified from their OD230/OD280 ratio, and the expected stoichiometry was confirmed by SDS-PAGE. The complex was assembled in hydrogenated GF buffer before being also purified by size-exclusion chromatography into the deuterated GF buffer. Prior to data collection, the Skp–uOMP complex was concentrated. Final D_2O concentrations were obtained by diluting with the appropriate D_2O and H_2O GF buffers. The resultant protein concentration was estimated from calculated extinction coefficients. The protein complexes were stored at 4 °C and used within 24 h after gel filtration.

SANS Data Collection

SANS measurements were performed on the NG3 30-m SANS instrument (Glinka, Barker et al., 1998) at the NIST Center for Neutron Research (NCNR) in Gaithersburg, Maryland. The neutron wavelength, λ, was 6 Å, with a wavelength spread, Δλ/λ, of 0.15. Scattered neutrons were detected with a 64 cm x 64 cm two-dimensional, position-sensitive detector with 128 x 128 pixels at a resolution of 0.5 cm $pixel^{-1}$. Data

reduction was accomplished using Igor Pro software (WaveMetrics, Lake Oswego, OR) with SANS macros developed at the NCNR (Kline, 2006). Raw counts were normalized to a common monitor count and then corrected for empty cell counts, ambient room background counts, and nonuniform detector response. Data were placed on an absolute scale by normalizing the scattering intensity to the incident beam flux for each individual pixel. Finally, the data were radially averaged to produce the scattering intensity I(q) to plot as I(q) versus q curves. Sample-to-detector distances of 5.0 m and 1.5 m were used in order to cover the range of 0.01 $Å^{-1} \leq q \leq 0.4$ $Å^{-1}$. The I(q) versus q scattering data obtained at the two instrument configurations were merged using the NCNR SANS reduction software (Kline, 2006). The I(q) versus q scattering data for the buffer was then subtracted from the data for the corresponding sample. The q-range covered by the data after buffer subtraction was dependent on the H_2O/D_2O ratio in the buffer.

Match Point and Contrast Variation Data Analysis

Determination of I(0) and R_g were performed using the Guinier approximation (Equations 2.6 and 2.7). The GNOM program (Semenyuk and Svergun, 1991) was used to determine the distance distribution function (Equation 2.8) to further confirm the R_g and I(0) values. Theoretical $\Delta\rho_1$ and $\Delta\rho_2$ values were calculated for the Skp and uOMP components and theoretical values were calculated for the complexes as a function of f_{D_2O} from the protein sequences using the Contrast Calculator (Sarachan, Curtis et al., 2013) module of the SASSIE (Curtis, Raghunandan et al., 2012) software assuming different percentages of deuteration for the OmpW and OmpA. The values were compared to those obtained from the Guinier analysis, allowing for experimental verification of the complex

match point and the amount of deuteration of the uOMP component in the measured samples. The match points of the individual components were then calculated from linear fits to $\Delta\rho$ versus f_{D2O} and $\Delta\rho^2$ versus f_{D_2O}. Both the Stuhrmann and parallel axis theorem analyses were performed using the program MULCh (Whitten, Cai et al., 2008). The contrasts and volumes were determined from the protein and buffer compositions and the radii of gyration and CM distances were determined using the Guinier R_g and I(0) values obtained from the data.

Molecular Dynamics Simulation of Skp-OmpA Complexes

The Skp homotrimer was modeled from the Skp crystal structure (PDB ID:1SG2; Korndorfer et al., 2004). The missing residues were included based on threefold symmetry considerations. The tentacles of the Skp homotrimer were then splayed out to specific separations using biased MD with NAMD (Phillips, Braun et al., 2005). A collective variable consisting of the distance between two groups defined as the α-carbon atoms of residues 35–80 in adjacent Skp monomers was defined. Then a harmonic potential with force constant 3.0 kcal $mol^{-}1$ $Å^{-2}$ was applied to the centers of the collective variable groups to bias the distance between adjacent tentacles to defined values. Simulations were done in vacuo using the CHARMM-22 force field parameters (MacKerell, Bashford et al., 1998). Stable distances were achieved in 10,000 steps of 1 fs. An ensemble of Skp homotrimer structures with separation distances of 10–119 Å at the tips of the helices and corresponding R_g values of 28.5–39.4 Å was generated.

Because the PP domain of OmpA folds as an independent unit (Danoff and Fleming, 2011) and is folded when it is in complex with Skp (Walton, Sandoval et al.,

2009), all-atom model structures of OmpA were constructed with an unfolded, but collapsed TM domain and natively folded PP domain as follows: a homology model of the *E. coli* PP domain (residues 187–316) was generated using Swiss-Model (Schwede, Kopp et al., 2003). An extended structure (θ = -78, φ = 149) containing the TM domain sequence of *E. coli* OmpA (residues 1–186) was built using PyMOL (The PyMOL Molecular Graphics System, version 1.7.4 Schrodinger, LLC, n.d.). This extended structure was partially collapsed using a torsion angle Monte Carlo procedure developed at Johns Hopkins University (REDUX, n.d.), resulting in an ensemble of structures with R_g values of 34–36 Å. These partially collapsed TM domain structures were manually connected to the above homology model of the PP domain. The combined TM–PP domain models were solvated with TIP3 water, neutralized with 0.2 M NaCl and subjected to standard MD equilibration using NAMD (Phillips, Braun et al., 2005) with the CHARMM-22 force field (MacKerell, Bashford et al., 1998). Successive rounds of equilibration with decreasing R_g potentials were carried out using the collective variables module in NAMD. Two of the resulting structures, one with an R_g of 42 Å and the other with 39 Å, were chosen for manual docking in PyMOL to the Skp models described above. Inspection of the docked structures suggested that the Skp model with R_g = 34 Å and tip separation of 46 Å fit the OmpA models best. Finally, two Skp–OmpA complex models with different orientations of the OmpA docked to Skp were solvated with TIP3 water, neutralized with 0.2 M NaCl and subjected to equilibration with standard MD using NAMD and the CHARMM-22 force field. The equilibrated complex model M1 contains Skp with R_g = 31.2 Å (S1) and OmpA with R_g = 43 Å and complex model M2 contains Skp with R_g = 32.5 Å (S2) and OmpA with R_g = 30 Å.

Modeling of the Skp–uOMP Complexes

Even though SANS is a low-resolution structural method, high-resolution structures derived from X-ray crystallography and NMR (along with MD and geometrical models) can be positioned to fit the SANS data in order to produce quasi-atomic structural models. Although the crystal structure of Skp is available, the bound uOMP TM domains remained unfolded when bound to Skp (Walton, Sandoval et al., 2009, Burmann, Wang et al., 2013). There are also no available data that link a specific region of Skp to a specific region of uOMP (Callon, Burmann et al., 2014). There is moreover intrinsic interplay of different regions of the uOMP on binding Skp. For example, in the case of PhoE, its N-terminal region is important for Skp binding, but its influence is strongly modulated by the OMP's C-terminal region (Harms, Koningstein et al., 2001). Therefore, OmpW and the TM portion of OmpA were modeled by a prolate ellipsoid of uniform scattering length density. The ellipsoid is the simplest model, i.e., a single geometric shape, that can be used to describe the uOMP TM domains, and it is adequate for the fitting of the SANS data, especially in the Guinier region. However, it should be noted that other options such as bead models, rigid body models consisting of multiple geometric shapes, and all-atom models can be more useful under certain circumstances. Such models are generally less symmetric than single geometric shapes and may better represent the overall shape of the biological molecule, especially at the shorter length scales on the order of 10 Å. This is especially important for fitting SAXS data, which usually have better signal to-noise than SANS data at q values beyond the Guinier range (Putnam, Hammel et al., 2007, Rambo and Tainer, 2010). All-atom models provide the

additional advantage of allowing modern MD methods to be employed to generate biologically relevant structures for comparison to SANS and SAXS data. For this reason, an all-atom model of Skp–OmpA was also tested against the SANS data, as is discussed later.

The prolate ellipsoid used for the TM domain of the uOMPs in both complexes was found by testing models of Skp and OmpW that best fit the Skp–OmpW contrast variation data. Three models were tested in which the axes were adjusted to create ellipsoids with R_g values of 21, 24, and 27 Å, (e21, e24, and e27, respectively) which match the range of possible R_g values found for bound OmpW, and volumes that essentially filled the Skp cavity, as suggested by both the I(q) and P(r) curves as a function of contrast. The ellipsoids were paired with Skp S1 and S2, from the M1 and M2 Skp–uOMP models described above and oriented such that the long axis was parallel to that of Skp. The CM distance between Skp and OmpW was varied for each model to correspond to a range calculated based on the R_g of Skp and the R_m value of the complex determined by the Stuhrmann and parallel axis theorem analyses. To reduce the R_g value of Skp to better match that obtained from the Skp–OmpW data, the S2 Skp structure was minimized for 1000 steps, subjected to a 10 ps MD simulation *in vacuo*, and then minimized again for 1000 steps using NAMD (Phillips, Braun et al., 2005), resulting in Skp model S3. To determine how well each model represented the experimental SANS data, SANS curves were calculated from each of the model structures using the programs SIMUL and SCAT (Hansen, 1990) as well as Xtal2sas (Heidorn and Trewhella, 1988, Krueger, Gorshkova et al., 1998). These curves were then compared to the Skp–OmpW SANS data to evaluate which Skp–OmpW model best fit the data. SIMUL was used to

build the OmpW ellipsoid and to populate it randomly with points of uniform scattering contrast, Δρ, corresponding to that of the deuterated OmpW and the D_2O composition of the buffer. Xtal2sas was used to read the atomic coordinates of the Skp structure (in PDB format) and to populate the structure randomly with points of scattering contrast that matched that of the amino acid residue in which the point was located. Care was taken to insure that the two volumes (Skp and OmpW) were filled with the same number density of points. SCAT was then used to calculate the scattering intensity of the complex, first by calculating P(r) as a function of r, using the contrast values associated with each scattering point, and then by performing a Fourier transform of P(r) to obtain I(q) as a function of q.

The calculated SANS curves were compared to the data at each contrast and goodness-of-fit was determined using the χ^2 equation

$$\chi^2 = \frac{1}{(N-1)} \sum_q \frac{\left(I_{exp}(q) - I_{calc}(q)\right)^2}{\sigma_{exp}(q)^2}, \qquad \text{Equation 2.14}$$

where $I_{exp}(q)$ is the experimentally determined SANS intensity curve, $I_{calc}(q)$ is the calculated intensity curve from the model structure and $\sigma_{exp}(q)$ is the q-dependent error of the $I_{exp}(q)$ values. The sum was taken over N = 60 data points. The goodness-of-fit to the entire contrast variation data set was determined by the average χ^2 value

$$\chi^2(\text{avg}) = \frac{1}{N} \sum_i \chi_i^2 \qquad \text{Equation 2.15}$$

where N is the number of contrast variation I(q) versus q scattering curves and χ_i^2 is the χ^2 value for the ith scattering curve. Once reasonable models for the Skp–OmpW complex were found, the Skp–OmpA complex was modeled using an ellipsoid of the same dimension for the TM portion of OmpA as found for OmpW. Guided by the structure of a closely related protein (PDB ID: 1R1M; Grizot & Buchanan, 2004), the PP

domain of OmpA was built using Swiss-Model (Schwede, Kopp et al., 2003) with ≈10 disordered residues of OmpA added at the N-terminus using the PSFGEN module of NAMD (Phillips, Braun et al., 2005) to act as a "tether." Significantly, this and other OmpA-like structures (PDB ID: 2K1S; Ramelot et al., unpublished) and (PDB ID: 2L26; Li et al., 2012) have similar three-dimensional folds and theoretical R_g values. Several models were tested in which the position of the TM domain of OmpA and the location at which the PP domain of OmpA was "tethered" approximately satisfied the R_g and CM distance parameters obtained from the Stuhrmann and parallel axis theorem analyses. The complex Monte Carlo module in SASSIE (Curtis, Raghunandan et al., 2012) was used to allow the disordered region of the PP domain of OmpA to vary to generate ensembles of structures from the starting Skp–OmpA structures for comparison to SANS data. Accepted (nonoverlapping) configurations were generated by sampling backbone dihedral angles using CHARMM-22 all-atom protein force field parameters (MacKerell, Bashford et al., 1998). The new configuration was checked for overlap of basis atoms, which were chosen as α-carbon atoms in this case. If the overlap distance between basis atoms was ≥ 3 Å, the new structure was accepted. Every 20th accepted structure was selected for further analysis in order to eliminate correlated dihedral angle moves.

A Skp–OmpA model with an all-atom representation of the TM domain of OmpA was also tested against the SANS data. Torsion angle MD (Chen, Won et al., 2005) was performed on the TM domain of OmpA (residues 1 to 171) from the M2 Skp–uOMP model described above in order to increase the R_g to match the value obtained for OmpW from the Skp–OmpW contrast variation data. This region was oriented such that the long axis was parallel to that of S3 Skp, with the C-terminal region near the top of the Skp

trimer. The PP domain of OmpA, as well as the disordered residues between the TM and PP domains and at the C-terminus, were then added using the PSFGEN module in NAMD (Phillips, Braun et al., 2005) to form the complete OmpA molecule (residues 1–326). The structure of the entire S3-OmpA complex was minimized using 1000 steps, subjected to a 10 ps MD simulation in vacuo to insure its stability, and then minimized again using 1000 steps, using NAMD (Phillips, Braun et al., 2005). The Complex Monte Carlo module in SASSIE (Curtis, Raghunandan et al., 2012) was used to explore possible conformations of the PP domain of OmpA as described above.

SANS curves were calculated for the ensembles of Skp–OmpA model structures as a function of contrast as described above for the Skp–OmpW models. In this case, Xtal2sas was also used to populate the PP domain of the OmpA structure (or both the PP and TM domains of OmpA in the all-atom model) with points of scattering contrast that matched that of the amino acid residue in which the point was located. Care was taken to insure that the volumes (Skp, TM-OmpA, and PP-OmpA) were filled with the same number density of points. Goodness-of-fit to the SANS data was determined using the same χ^2 relations as above.

Examination of a plot of χ^2 versus R_g at each individual contrast, or χ^2(avg) versus R_g(avg) for the entire contrast variation data set, provides an idea of how well the individual structures generated from each starting structure fit the data as well as which starting structure produces the overall best fit to the data. The best- (lowest χ^2(avg)) and worst-fit (highest χ^2(avg)) model SANS curves were noted for each case, along with the average model SANS curve from the entire ensemble of accepted structures. These curves were plotted along with the experimental SANS data to aid in the visualization of

the quality of the fits to the data. Surface density plots representing the total configuration space examined by all of the accepted structures were generated and compared to that representing the best-fit family of structures in each case. The best-fit family of structures was chosen based on those giving the lowest ≈10% of the χ^2(avg) values for each series. This cutoff was chosen arbitrarily based on the shape of the χ^2(avg) versus R_g(avg) curves and used to analyze the results further.

2.3 Results

SANS Data

The SANS data for Skp alone in 0% and 98% D_2O buffers are shown in figure 2.2. Guinier analysis of the data resulted in R_g values of 34 ± 2 Å in 0% D_2O buffer and 32.9 ± 0.3 Å in 98% D_2O buffer. The Guinier-derived I(0) values of 0.125 ± 0.005 cm^{-1} in 0% D_2O buffer and 0.160 ± 0.002 cm^{-1} in 98% D_2O buffer agree well with the calculated values of 0.14 and 0.15 cm^{-1}, respectively, using equations 2.4 and 2.5 along with the M_w calculated from the Skp sequence. This verifies that Skp is a trimer in solution, in agreement with the crystal structures (Korndörfer, Dommel et al., 2004, Walton and Sousa, 2004). Four SANS data sets were collected on both the Skp–OmpW and Skp–OmpA complexes. The Skp–OmpW data were collected in buffers with 0%, 30%, 80%, and 98% D_2O while the Skp–OmpA data were collected in buffers with 0%, 15%, 30%, and 98% D_2O, as shown in figure 2.3, along with their corresponding distance distribution functions, P(r) versus r. The R_g and I(0) values obtained from both Guinier and P(r) analyses are listed in tables 2.1 and 2.2 for Skp–OmpW and Skp–OmpA, respectively.

Match Point Analysis and Quantification of Omp Deuteration

$\sqrt{I(0)/c}$ versus f_{D_2O} plots using the Guinier-derived values in tables 2.1 and 2 are shown for the Skp–OmpW and Skp–OmpA complexes in figure 2.4, along with the calculated curves assuming different percentages of deuteration for the uOMP component (30%, 40%, and 50% deuteration for OmpW and 40%, 50%, and 60% deuteration for OmpA). The match point for the Skp–OmpW complex was found to be 51% ± 2% D_2O, while that for the Skp–OmpA complex was found to be 57% D_2O ± 2% D_2O. It can be seen from figure 2.4 A that the values for Skp–OmpW as a function of f_{D_2O} fall closer to the curve for 30% deuteration of OmpW. This value seems low given that OmpW was grown in a medium with a starting value of 60% D_2O. The data point for 0% D_2O buffer does match that for 50% deuterated OmpW. On the other hand, the values for Skp–OmpA as a function of f_{D_2O} fall mainly on the curve for 50% deuterated OmpA, which seems more reasonable, given the starting value of 60% D_2O in the growth medium. The data point for 98% D_2O agrees better with the 40% deuterated OmpA curve. It is possible that the concentration estimated by extinction coefficients is higher than the actual concentration of the complex during the SANS measurement, resulting in a value that is low for this data point. Given the results shown in figure 2.4B, the amount of deuteration for OmpA was estimated at 50% ± 5%. There is no compelling reason why the amount of deuteration should be different in OmpW and OmpA, as they were prepared under the same conditions. Because using either 30% or 50% for the amount of deuteration for OmpW did not change the subsequent Stuhrmann and parallel axis theorem analyses results significantly, 50% ± 5% deuteration was also estimated for OmpW. Given this

value for both OmpW and OmpA, the match points of the Skp–OmpW and Skp–OmpA complexes were calculated to be 55% D_2O ± 2% $D_2O$and 60% D_2O ± 2% D_2O, respectively. Additionally, the match points of the components were determined to be 44% D_2O for Skp and 80% D_2O ± 7% D_2O for OmpW and OmpA.

The results in figure 2.4A and B indicate that the Skp–uOMP complexes were present in all samples analyzed and were formed by three Skp monomers bound to a single uOMP. This stoichiometric ratio is consistent with previous binding data (Qu, Mayer et al., 2007, Moon, Zaccai et al., 2013). It is important to emphasize that no urea was present, and without Skp, the uOMPs would have precipitated over the course of the SANS data collection.

Stuhrmann and Parallel Axis Theorem Analyses

Both Stuhrmann and parallel axis theorem analyses were performed on the contrast variation SANS data and the results are listed in tables 2.3 and 4, respectively. 50% deuteration was assumed for both OmpW and OmpA, with the $\Delta\rho$ values calculated accordingly for each contrast. R_g changes very little as a function of contrast for the Skp–OmpW complex. Therefore, R_g^2 versus $\Delta\rho^{-1}$ is nearly flat and the fit to equation 2.10 results in large uncertainties for both α and β. A similar argument can be made for fits to equation 2.11. As a result, these analyses only allow determination of R_g for the Skp component of the complex, while R_g for the OmpW component and the CM distance, D_{CM}, have large uncertainties. Importantly, R_m, the radius of gyration of an equivalent complex of uniform scattering length density, is also well determined from the Stuhrmann analysis (Equation 2.10). Thus, three possible R_g (OmpW) versus D_{CM} curves

were calculated from Equation 2.10 using the well-determined values obtained for R_g (Skp) and R_m (along with the values obtained by adding and subtracting their corresponding errors) and the results are shown in figure 2.5. This information was used to constrain the size of OmpW and the CM distance between Skp and OmpW in the Skp–OmpW model structures as described below.

Conversely, the change in R_g as a function of contrast is more pronounced for the Skp–OmpA complex. Thus, R_g for both Skp and OmpA, as well as D_{CM}, were well determined in this case and these values were used directly to build the starting model structures for the Skp–OmpA complex.

Skp Structure in Solution

Figure 2.6A shows the starting model for Skp based on the X-ray crystal structure as described in Section 3. The calculated R_g based on the atomic coordinates is 30.6 Å. This is about 3 Å smaller than the Guinier-derived R_g. The model structure obtained after the R_g-constrained MD simulation, shown in figure 2.6B, has a calculated R_g of 33.5 Å, in good agreement with the Guinier derived R_g value. Importantly, the Skp β-strand "body" through which oligomerization occurs did not alter in conformation between the two models. This region merely served as a "hinge" point for the opening of the Skp helical tentacles. This structural pivot was also identified from NMR data collected from the OmpX–Skp complex (Burmann, Wang et al., 2013).

Skp–OmpW Structure in Solution

The fact that R_g does not change appreciably with contrast for the Skp–OmpW complex means that the R_g values of the two components are close to each other. Thus, a number of possible values for R_g (OmpW) and CM distance, D_{CM}, in figure 2.5 can be ruled out, given the R_g (Skp) and R_m values from the Stuhrmann analysis in table 2.3. Several model structures were tested with Skp structures S1, S2, and S3, with R_g values of 31.2, 32.5, and 31.6 Å, respectively, OmpW represented by e21, e24, and e27 ellipsoids with R_g values of 21, 24, and 27 Å , respectively, and D_{CM} in the range of 25–30 Å. Representative model structures tested are pictured in figure 2.7 and are listed in table 2.5, along with the R_g values of the individual components and the CM distance between

them. table 2.1 shows that R_g is largest for the Skp–OmpW complex in 0% D_2O buffer and it is about 1 Å smaller in 98% D2O. A comparison of these results with those in table 2.6, which lists R_g as a function of contrast for the model structures, shows that only model S1e27 satisfies this requirement. However, R_g in 80% D2O is slightly low in this model. This means that the R_g value for Skp is slightly too small, because Skp dominates the scattering at this contrast. Furthermore, the S1 Skp structure is too symmetric, as can be seen in figure 2.8A, which shows the model SANS curves for S1, S2, and S3 Skp. The curve for S1 Skp (shown in the S1e27 complex in figure 2.7) shows a pronounced peak at $q = 0.15$ Å^{-1} that is not observed in the data, which is due to the threefold symmetry of the three tentacles. Even the S2 Skp model (shown in the S2e21 and S2e27 complexes in figure 2.7) results in a SANS curve that shows some structure at this q value, which was reduced substantially by performing a 10 ps MD simulation *in vacuo*, as described in

Section 3, to create the less symmetric S3 Skp (shown in the remaining complexes in figure 2.7). In addition, the model SANS curve for an ellipsoid with R_g = 27 Å features a strong peak at q = 0.18 $Å^{-1}$, since ellipsoids are highly symmetric structures. The peak is less pronounced for ellipsoids of smaller R_g, as shown in figure 2.8B. Given these results, model structures with S1 Skp and e27 ellipsoids do not represent the data well. Model S2e21, consisting of S2 Skp and the smaller e21 ellipsoid with R_g = 21 Å , also is not a good representation of the data; the R_g is larger in 98% D_2O than in 0% D_2O and R_g in 30% D2O is too small, indicating that the e21 ellipsoid is too small, because OmpW dominates the scattering at this contrast. On the other hand, a reasonable working model of the structure of Skp–OmpW in solution was obtained using S3 Skp and the e24 ellipsoid with R_g = 24 Å, with the two components positioned such that DCM = 30 Å to satisfy the blue (dark gray in the print version) middle curve in figure 2.5 in which R_g(Skp) = 31.6 Å and R_m = 32.3 Å. While R_g in 98% D_2O is not smaller than that in 0% D_2O, the two values are equal and the discrepancy may be due to the CM distance, D_{CM}, being slightly too large. Several views of S3e24 are shown in figure 2.9, which also shows the location of the CM for both the Skp and OmpW components.

Model SANS curves from S3e24 are shown along with the data in figure 2.10. They agree well with the data for all contrasts below q = 0.1 $Å^{-1}$, beyond which the data for 30% D_2O and 80% D_2O become too noisy for comparison. The model SANS curves for 0% D_2O and 30% D_2O show a peak at q = 0.18 $Å^{-1}$ that is not seen in the data due to the symmetry of the ellipsoid model for OmpW.

Skp–OmpA Structure in Solution

Starting model structures for Skp–OmpA were constructed based on the working model for Skp–OmpW above, incorporating the e24 ellipsoid to represent the TM domain of OmpA. The ellipsoid was positioned in the same location as for the Skp–OmpW model and also higher in the Skp cavity for comparison. The PP domain of OmpA was tethered near the center of the Skp cavity in slightly different positions as described in Section 3 such that the starting values of R_g (OmpA), including both the TM and PP domains, and D_{CM} approximately agreed with the values obtained in tables 2.3 and 4. The starting values of these parameters did not need to be exact because the PP domain would be repositioned during the SASSIE runs to find the best-fit structures to the data. Representative starting models for Skp–OmpA are shown in figure 2.11, including an all-atom model, in which the residues at the N-terminus of the PP domain are located high in the Skp cavity. The starting R_g values are shown as a function of contrast in table 2.7. Model S3e24PP1 was not pursued further because the R_g values at all contrasts are too small and the disordered residues that tether the PP domain are shorter than in the other models. Thus, structures with larger R_g values were unlikely to be generated during the SASSIE runs. Short SASSIE runs producing a few hundred structures were performed on the rest of the starting model structures. Models S3e24PP2 and S3e24PP3 were found to be subsets of S3e24PP4 and S3e24PP5, respectively. Thus, longer SASSIE runs were only performed on models S3e24PP4 and S3e24PP5.

SASSIE runs performed on models S3e24PP4, S3e24PP5, and the all-atom model, S3-OmpA, produced 1650 accepted structures for S3e24PP4 and S3e24PP5 and 2144 accepted structures for S3-OmpA. $\chi 2$ versus R_g plots are shown as a function of contrast

for models S3e24PP4 and S3e24PP5 in figure 2.12. The larger χ^2 values observed for the model SANS curves in 98% D2O indicate that the S3 model for Skp does not describe the data as well for Skp–OmpA as it does for Skp–OmpW. Although the match point of the uOMPs is around 80% D_2O, the scattering from this component is still relatively weak in 98% D_2O compared to that of the Skp component. Thus, the differences in the 98% D_2O data can be ascribed mainly to the structure of Skp. The 98% D_2O data for Skp–OmpA (figure 2.3B) shows a distinct peak at $q \approx 0.18$ Å^{-1}, where as the data at the same contrast for Skp–OmpW (figure 2.3A) lacks this peak. Thus, the structure of Skp bound to OmpA is more symmetric with respect to the location and shape of the three tentacles. This leads to the conclusion that a model more like S1 Skp, but with a larger R_g value more like S3 Skp, would result in a better fit of the Skp–OmpA models to the 98% D_2O data.

The χ^2(avg) versus R_g(avg) plots for both the S3e24PP4 and S3e24PP5 model structures are shown in figure 2.13. Better χ^2(avg) values are obtained for R_g(avg) > 34 Å for model S3e24PP5, although there are not as many structures with R_g(avg) > 38 Å, represented by the rectangular region, which denotes the lowest ≈10% of χ^2(avg) values. The χ^2(avg) values in this best-fit region are lower for S3e24PP5, suggesting that the TM domain of OmpA sits higher in the Skp cavity in the Skp–OmpA complex than OmpW inthe Skp–OmpW complex. χ^2 versus R_g plots are shown as a function of contrast for the all-atom Skp–OmpA model structure in figure 2.14. A comparison to figure 2.12 reveals the same high χ^2 values in 98% D_2O, as expected because the S3 model for Skp was used in both cases. On the other hand, the χ^2 values overallare smaller at all contrasts compared to the S3e24PP5 model. The same is true for the χ^2(avg) versus R_g(avg) plot

for Skp–OmpA in figure 2.15, which shows lower χ^2(avg) values overall in the best-fit region where R_g(avg) is between 37 and 41 Å. While this model shows only one of many possible configurations for the TM domain of OmpA, it does provide an example ofan alternate tethering strategy, and thus a different region of conformation space, for the PP domain of OmpA than is represented by S3e24PP4 or S3e24PP5.

Plots of the data at each contrast compared to the calculated S3e24PP5 SANS curves for the best- and worst-fit structures, as well as the average SANS curve for the entire ensemble, are shown in figure 2.16. As explained earlier, the discrepancy between the calculated SANS curves and the SANS data in 98% D_2O is due to the use of the S3 model for Skp for both Skp–OmpW and Skp–OmpA. The portion of the Skp–OmpA data near not exist in the Skp–OmpW data. Additional MD simulations were not performed on Skp to try to find Skp–OmpA models that better fit the 98% D2O data, since this would have only resulted in a shift in χ^2(avg) by a constant value in figure 2.13 and 2.15. Figure 2.16 shows that the average SANS curve for the entire ensemble of S3e24PP5 structures fits the data almost as well as that for the best-fit single structure in each case, indicating that the entire generated ensemble of structures is also a reasonable fit to the SANS data. Similar plots are shown for the calculated S3-OmpA SANS curves compared to the data at each contrast in figure 2.17. Again, the average curve for the entire ensemble fits the data almost as well as that for the best-fit single structure in each case.

The calculated SANS curves from the global best-fit individual structure are shown in figure 2.18 at each contrast for S3e24PP5 and in figure 2.19 for S3-OmpA, along with the corresponding global best- and worst-fit structures for comparison. The resultant R_g values as a function of contrast for the global best-fit individual structures are

listed in table 2.8 and the R_g values for the Skp and OmpA components, as well as the CM distance, D_{CM}, for the global best-fit individual structures are listed in table 2.9. In agreement with the χ^2(avg) versus R_g(avg) curves, the parameters obtained from S3e24PP5 are in better agreement with the data than those obtained from S3e24PP4, although the D_{CM} value is still a little large compared to those obtained from the data. The parameters obtained from S3-OmpA are in very good agreement with those obtained from the data. Surface density plots for S3e24PP5 and S3-OmpA showing the configuration space covered by the PP domain for the ensemble (gray (light gray in the print version)) and the best-fit ensemble (green(medium gray in the print version)) are presented in figure 2.20. Skp is shown in black and the TM domain of OmpA is in red (dark gray in the print version). These plots show that the PP domain can assume multiple positions and still produce structures that fit the data well. The PP domain in S3e24PP5 was tethered such that it could occupy all positions shown in gray (light gray in the print version) and those structures that fall in the rectangle in the χ^2(avg) versus R_g(avg) plot (figure 2.13) can take the positions shown in green (medium gray in the print version). The gray (light gray in the print version) areas that do not overlap with the green (medium gray in the print version) are as represent the conformations that do not fall in the rectangle. The PP domain was tethered higher within the Skp cavity in S3-OmpA, so a different part of conformation space was explored using SASSIE. But, the result is similar to that of S3e24PP5 in that there is only a small region of conformation space (the gray (light gray in the print version) that is not overlapping with the green (medium gray in the print version)) that represents the conformations that do not fall in the rectangle in the χ^2(avg) versus R_g(avg) plot (figure 2.15). Both surface density plots reveal that there

is a portion of conformation space around Skp that is not sampled by the PP domain of OmpA.

2.4 Discussion

Skp Alone in Solution

The SANS data for Skp alone in solution are consistent with a structure similar to that of the Skp X-ray crystal structure (Walton, Sandoval et al., 2009), but with the helical tentacles more open resulting in a slightly larger R_g value than that for the crystal structure (figure 2.6). This splaying of Skp is not unexpected because the tips of the helical tentacles are positively charged and flexible, based on thefact that two of them are missing in the X-ray crystal structure (PDB ID:1U2M). Also, recent NMR data from Skp–uOMP complexes demonstrated that residues 89–93, higher up the Skp tentacles closer to the body, display increased helicity on substrate binding and could therefore be at the pivot position for the splaying of the tentacles (Burmann, Wang et al., 2013). A crystal structure of the Tim 9–10 complex (PDB ID: 2BSK) also lacks density for the tentacles, possibly due to inherent flexibility (Webb, Gorman et al., 2006, Ryan, & Gulbis, 2006, Beverly, Sawaya et al., 2008).

Skp–uOMP Structure and Binding Mechanism

Comparison of the SANS data for Skp alone in 98% D_2O and for the Skp–OmpA complex in 98% D_2O confirm that the conformation of Skp when bound to OmpA is similar to that when it is alone in solution, although it has a smaller R_g when bound to OmpA. When bound to OmpW, Skp adopts a conformation with a less symmetric

configuration of the three tentacles and with an even smaller R_g. However, the R_g value for Skp in both complexes remains larger than that observed for the crystal structure. These data suggest that, in the absence of substrate, the Skp tentacles are in an open conformation (figure 2.6B); on binding an uOMP, Skp collapses in a clamp-like mechanism towards the more closed conformation represented by the crystal structure (figure 2.6A). This flexibility of the Skp tentacles may allow for different conformations of the Skp α-helices to accommodate different uOMPs with TM domains larger or smaller than those of OmpW and OmpA. The SANS data revealed that the OmpA and OmpW TM domains are unfolded, but constrained in size, when bound to Skp. The R_g value of theTM region ($R_g \approx 24$ Å) is significantly larger in comparison to that calculated from the crystal structure of the folded TM domain ($R_g \approx 14$ Å) (Hong, Patel et al., 2006) (Pautsch and Schulz, 1998). A similar radius (21 Å) was measured by NMR for paramagnetic spin-labeled OmpX that was bound to Skp (Burmann, Wang et al., 2013), in which case the volume occupied by OmpX was found to contain ≈50% water. Like OmpW, OmpX (16.5 kDa) folds into an eight-stranded β-barrel integral membrane protein (Vogt and Schulz, 1999). Significantly, if these TM domains adopted a random coil state, as when fully chemically denatured, the theoretical R_g would be much larger (Fleming, Fitzkee et al., 2005). For example, the radius of OmpX in 8 M urea was determined by NMR to be 45 Å (Burmann and Hiller, 2012). Additionally, an R_g of the unfolded TM domain of OmpA in solution under folding conditions can be estimated from the sedimentation coefficient to be 32 Å (Danoff and Fleming, 2011). This value is still considerably larger than the value determined by SANS, which means that Skp sequesters the unfolded TM domains of the uOMPs in a conformation that is smaller than

they would naturally adopt in solution.

The SANS analysis of Skp–OmpA and Skp–OmpW complexes suggests a model for the binding of the uOMP TM domains near the tips of the Skp tentacles (S3e24 in figures 2.7 and 2.9; S3e24PP5 and S3-OmpA in figure 2.11). This region of Skp, which is also involved in the initial capture of OmpA (Lyu, Shao et al., 2012 & Zhao, 2012), is highly positively charged due to a large number of arginine residues. Conversely, the uOMP TM domains are typified by their overall negative charge with a pI ranging between 4 and 7 (Jarchow, Luck et al., 2008). The Skp–uOMP complex may consequently be stabilized by continued electrostatic interactions between the tentacle tips and the uOMP. Such a notion is supported by fluorescence studies that show weakening of observed binding in the presence of high salt (Qu, Mayer et al., 2007). A similar arrangement of the TM domain of OmpX was found from TROSY NMR data collected from a SurA–OmpX complex. The results support the notion that the OmpX TM domain has the same preferred conformational ensemble either in complex with Skp or with SurA (Burmann and Hiller, 2012). The similarity between uOMP spectra may be explained by the small surface area by which both these chaperones interact with uOMPs. In particular, SurA can bind short peptides (7–14 residues long) with a specific tripeptide motif (Bitto and McKay, 2003). Additionally, amino acid deletion experiments on the small Tim complexes (Vergnolle, Baud et al., 2005) and prefoldin (Siegert, Leroux et al., 2000, Hartl, & Moarefi, 2000) show that the tips of their tentacles are required for substrate binding. The SANS-derived structures show the OmpW and OmpA TM domains are bound to Skp in a position analogous to that of actin bound to prefoldin (Lundin, Leroux et al., 2010). However, only the tips of the prefoldin tentacles interact

with the unfolded actin and no electron density can be observed in the prefoldin cavity.

The model structures for Skp–OmpW and Skp–OmpA fit the SANS data best when the TM domain of OmpA is centered farther from the tips of the Skp tentacles, and thus closer to the body of Skp, than that of OmpW. In both cases, the closed conformation of Skp and the size and relative positions of the TM domains imply that a portion of the uOMP substrate is protruding outside the Skp cavity. This observation is consistent with interaction data from NMR analysis of Skp–uOMP complexes (Walton, Sandoval et al., 2009, Burmann, Wang et al., 2013, Callon, Burmann et al., 2014), which demonstrated that binding is focused on the Skp tentacles, but there is no single preferred conformation of uOMP in this region of Skp. Because NMR only maps the presence of interactions, any substrate residues that extend away from the Skp cavity would be spectroscopically silent. Thus, the SANS-derived structure models implying that individual strands of uOMP migrate out of the Skp cavity are consistent with the NMR results. The combination of the inherent flexibility in the Skp tentacles and the lack of a specific structure for the bound uOMPs apparently insure that the uOMPs are sufficiently shielded from self-aggregation until the outer membrane is reached even if portions of the unfolded region protrude outside of the Skp cavity. The best-fit structures to the Skp–OmpA SANS contrast variation data set using models with both ellipsoidal and all-atom representations of the TM region support a model for an elongated Skp–OmpA complex with a folded PP domain of OmpA protruding out near the bottom of the Skp tentacles. However, the ensembles of structures that represent the lowest ≈10% of χ^2(avg) values include those in which the PP domain of OmpA exits out to the side of Skp, and its location in any individual structure is not necessarily fixed. Furthermore, the calculated

average SANS curve for the entire generated ensemble of structures fits the data almost as well as that from the best-fit structure. Thus, the solution can consist of a similar ensemble of structures in which the PP domain of OmpA assumes a number of different positions outside the Skp cavity and still be consistent with the SANS data.

Release Mechanism

The SANS data show that the Skp tentacles change conformation to accommodate different uOMP TM domains using a simple clamp-like mechanism used by jellyfish-like chaperones (Stirling, Bakhoum et al., 2006). However, the fact that a portion of the uOMP polypeptide resides outside the Skp cavity implies that there is no requirement for a large-scale rearrangement of Skp for uOMP to be exchanged with an empty Skp trimer (Burmann, Wang et al., 2013) and other molecular chaperones (Sklar, Wu et al., 2007, Schwalm, Mahoney et al., 2013), or for uOMP to be presented to the Bam complex or to the bacterial outer membrane. While the PP domain of OmpA can exit from openings at either the base or the sides of the tentacles, there is an ample fraction of conformation space not occupied by the PP domain for any individual Skp–OmpA structure. Thus, the SANS data do not rule out the hypothesis that OmpA would exit the chaperone from its side. This is consistent with the capability of Skp to directly interact with membranes (De Cock, Schäfer et al., 1999), as well the idea that Skp may facilitate the folding of OmpA into negatively charged lipids or lipopolysaccharides (Bulieris, Behrens et al., 2003, Patel, Behrens-Kneip et al., 2009, Patel and Kleinschmidt, 2013). Even if the presence of the negatively charged uOMP counteracts the positive charges at the base of the Skp tentacles, a large section of the chaperone would still be available to bind lipids. In

particular, in the SANS model of the Skp–OmpA complex, the three symmetrically related, highly conserved regions of Skp that are centered on residues E29, K77, and R88 (Burmann, Holdbrook et al., 2015, Bond, & Hiller, 2015) in the middle of each tentacle (Walton and Sousa, 2004) would remain accessible, and only one binding site would be lost during those times when the PP domain occupied that portion of conformation space.

Conclusions

Deuterium labeling of the uOMP component in Skp–OmpW and Skp–OmpA complexes coupled with SANS and contrast variation has enabled the measurement of individual R_g values of Skp and uOMP when in complex with each other. This experimental strategy also provided information on the distances between their centers of mass. These unique structural properties cannot be obtained using unlabeled complexes or other methodologies. Simultaneous analysis of the data allowed a determination of structure models consistent with the entire contrast variation data set. In conjunction with independent studies using other techniques, the SANS data were used to postulate how uOMPs are captured by Skp and subsequently released to the outer membrane.

Using the R_g information for Skp and the contrast variation data to constrain the model structures, it was determined that Skp can undergo conformational rearrangement to accommodate its client: Skp alone in solution is larger than Skp bound to OmpA, which is larger than Skp bound to OmpW, which is larger than Skp in protein crystals. Furthermore, the Skp tentacles are less ordered with respect to each other in the Skp–OmpW complex than in the Skp–OmpA complex. These findings would not have been revealed without deuterium labeling that allowed the separation of the Skp structure as it

exists within the Skp–uOMP complexes. Skp dominated the scattering profile in 98% D_2O buffer, where differences were clearly observed between the Skp–OmpW data and the Skp–OmpA data. These differences were not readily observed in the 0% D_2O data, for instance, where Skp does not contribute as much to the scattering. This result suggests that the Skp tentacles have the ability to adjust to accommodate different uOMPs. Accordingly, because the uOMPs were labeled and distinguishable from the Skp component, it was possible to determine that the OmpW TM domain likely sits lower in the Skp cavity than the OmpA TM domain. Finally, the resultant structure density plots mapping the conformation space of the OmpA PP domain revealed that it could exit from openings at either the base or the sides of the Skp tentacles. However, there is an ample fraction of conformation space not occupied by the PP domain for any individual Skp–OmpA structure. Given the flexibility of the Skp tentacles and the fact that a portion of the TM domain of uOMP resides outside the Skp cavity, it can be postulated that a low-energy path will always exist such that the TM domains of both OmpW and OmpA can be delivered to the outer membrane or other chaperones. Specifically, the SANS data do not rule out the premise that the TM domain can exit from the side of Skp, even when a PP domain is present as in the case of OmpA.

2.5 Tables and Figures

Table 2.1. Skp-OmpW Parameters

	0% D_2O	30% D_2O	80% D_2O	98% D_2O
Concentration[a] (mg mL^{-1})	3.45	3.8	1.14	1.25
Guinier R_g (Å)	32.7 ± 0.2	31 ± 1	31.5 ± 0.4	31.6 ± 0.2
Guinier $I(0)$ (cm^{-1})	0.223 ± 0.001	0.027 ± 0.001	0.020 ± 0.001	0.063 ± 0.001
R_g from $P(r)$ (Å)	32.9 ± 0.1	29.0 ± 0.2	31.32 ± 0.03	31.3 ± 0.2
$I(0)$ from $P(r)$ (cm^{-1})	0.240 ± 0.001	0.025 ± 0.001	0.020 ± 0.001	0.063 ± 0.001
D_{max} (Å)	105 ± 5	83.5 ± 1.5	89 ± 4	93.5 ± 1.5

[a]Errors on the concentration values are about 5%.
Errors on R_g, $I(0)$, and D_{max} represent the standard error of the mean for three different Guinier or $P(r)$ fits.

Table 2.2. Skp-OmpA Parameters

	0% D_2O	15% D_2O	30% D_2O	98% D_2O
Concentration[a] (mg mL^{-1})	5.6	5.4	5.3	1.31
Guinier R_g (Å)	39.7 ± 0.1	39.8 ± 0.3	39.1 ± 0.3	36.3 ± 0.7
Guinier $I(0)$ (cm^{-1})	0.462 ± 0.002	0.261 ± 0.004	0.119 ± 0.001	0.056 ± 0.001
R_g from $P(r)$ (Å)	41.1 ± 0.4	42.0 ± 0.8	41.8 ± 0.3	38.1 ± 0.1
$I(0)$ from $P(r)$ (cm^{-1})	0.475 ± 0.003	0.263 ± 0.002	0.120 ± 0.001	0.056 ± 0.001
D_{max} (Å)	135 ± 5	137 ± 7	133 ± 3	127 ± 3

[a]Errors on the concentration values are about 5%.
Errors on R_g, $I(0)$, and D_{max} represent the standard error of the mean for three different Guinier or $P(r)$ fits.

Table 2.3. Stuhrmann Analysis for Skp-uOMP Complexes

	Skp–OmpW	Skp–OmpA
α	64 ± 55	242 ± 34
β	41 ± 162	461 ± 66
R_m (Å)	32.3 ± 0.5	39.3 ± 0.2
χ^2	4.75	0.4
R_g (Skp) (Å)	31.5 ± 0.5	33.5 ± 0.8
R_g (Omp) (Å)	32 ± 6	36.1 ± 0.8
D_{CM} (Å)	13 ± 26	38 ± 3

50% deuteration was assumed for OmpW and OmpA. χ^2 is from the fit to Eq. (10).

Table 2.4. Parallel Axis Theorem Analysis for Skp-uOMP Complexes

	Skp–OmpW	Skp–OmpA
R_g (Skp) (Å)	31.3 ± 0.6	32.7 ± 0.9
R_g (Omp) (Å)	32 ± 6	35.7 ± 0.9
D_{CM} (Å)	14 ± 28	40 ± 3

50% deuteration was assumed for OmpW and OmpA.

Table 2.5. Skp-OmpW Model Structures

Model	R_g (Skp) (Å)	R_g (OmpW) (Å)	R_g (Complex) (Å)	D_{CM} (Å)
S2e21	32.5	21.2	32.7	30.5
S1e27	31.2	27.6	32.1	28.6
S2e27	32.5	27.6	32.1	28.6
S3e27s	31.6	27.6	31.9	25.0
S3e27	31.6	27.6	32.1	29.2
S3e24	31.6	23.9	32.0	30.9

Values are from the coordinates only. Contrast was not taken into account.

Table 2.6. Skp-OmpW Model Structure R_g as a Function of Contrast

Model	0% D_2O R_g (Å)	30% D_2O R_g (Å)	80% D_2O R_g (Å)	98% D_2O R_g (Å)
S2e21	31.7	27.8	32.4	32.8
S1e27	32.3	30.9	30.4	31.6
S2e27	33.1	30.1	32.4	33.3
S3e27s	32.4	31.2	31.9	32.4
S3e27	33.3	31.5	31.9	33.0
S3e24	32.4	29.8	31.9	32.6
Guinier	32.7 ± 0.2	31 ± 1	31.5 ± 0.4	31.6 ± 0.2
$P(r)$	32.9 ± 0.1	29.0 ± 0.2	31.32 ± 0.03	31.3 ± 0.2

Table 2.7. Skp-OmpA Starting Model Structure R_g as a Function of Contrast

Model	0% D_2O R_g (Å)	30% D_2O R_g (Å)	98% D_2O R_g (Å)
S3e24PP1	37.4	35.7	35.9
S3e24PP2	40.3	38.9	37.2
S3e24PP3	40.1	37.6	37.7
S3e24PP4	42.1	40.3	38.3
S3e24PP5	42.3	41.4	37.9
S3-OmpA	39.1	37.5	36.3
Guinier	39.7 ± 0.1	39.1 ± 0.3	36.3 ± 0.7
$P(r)$	41.1 ± 0.4	41.8 ± 0.3	38.1 ± 0.1

Table 2.8. Skp-OmpA Global Best-Fit Structure R_g as a Function of Contrast

SASSIE Run[a]	Structure Number[b]	0% D_2O R_g (Å)	15% D_2O R_g (Å)	30% D_2O R_g (Å)	98% D_2O R_g (Å)
S3e24PP4 (39.9)	442 (1650)	41.4	40.8	39.2	38.3
S3e24PP5 (40.8)	740 (1650)	42.2	41.8	41.1	38.2
S3-OmpA (38.7)	1188 (2144)	39.9	39.7	38.9	36.4
Guinier		39.7 ± 0.1	39.8 ± 0.3	39.1 ± 0.3	36.3 ± 0.7
$P(r)$		41.1 ± 0.4	42.0 ± 0.8	41.8 ± 0.3	38.1 ± 0.1

[a]Average R_g (Å) for the global best-fit structure is in parenthesis.
[b]Total number of structures is in parenthesis.

Table 2.9. Component R_g values for Skp-OmpA Global Best-Fit Structure

SASSIE Run	Structure Number	R_g (Skp) (Å)	R_g (OmpA) (Å)	D_{CM} (Å)
S3e24PP4	442	32.0	33.8	51.3
S3e24PP5	740	32.0	37.8	46.6
S3-OmpA	1188	31.7	36.6	40.5
Stuhrmann		33.5 ± 0.8	36.1 ± 0.8	38 ± 3
Parallel axis		32.7 ± 0.9	35.7 ± 0.8	40 ± 3

Values are from the coordinates only. Contrast was not taken into account.

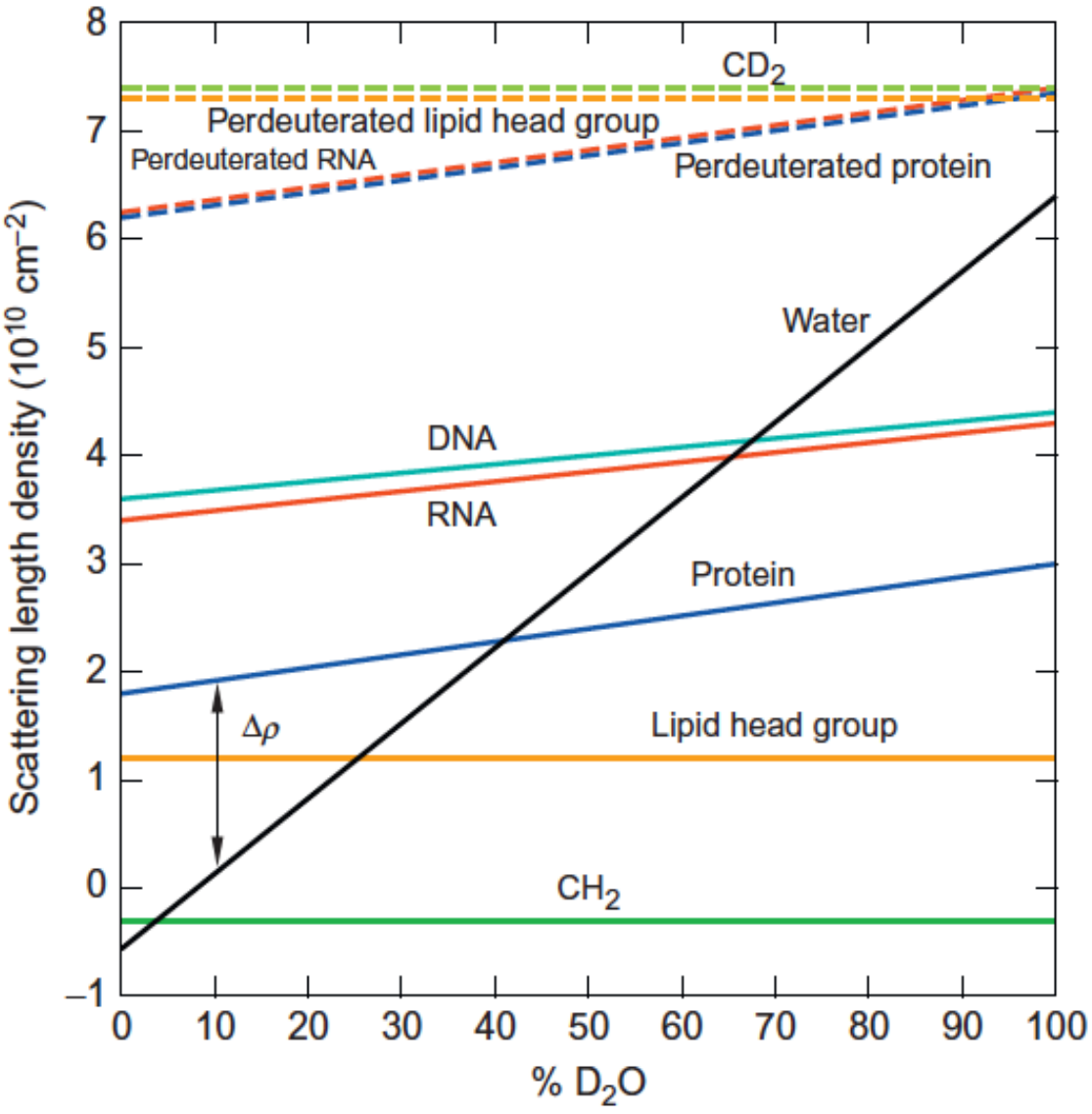


Figure 2.1. Plot of the neutron scattering length density versus % D_2O in the solvent for water compared to those for protein, RNA, DNA, and the components of lipids (lipid head group and CH_2), along with their perdeuterated counterparts.

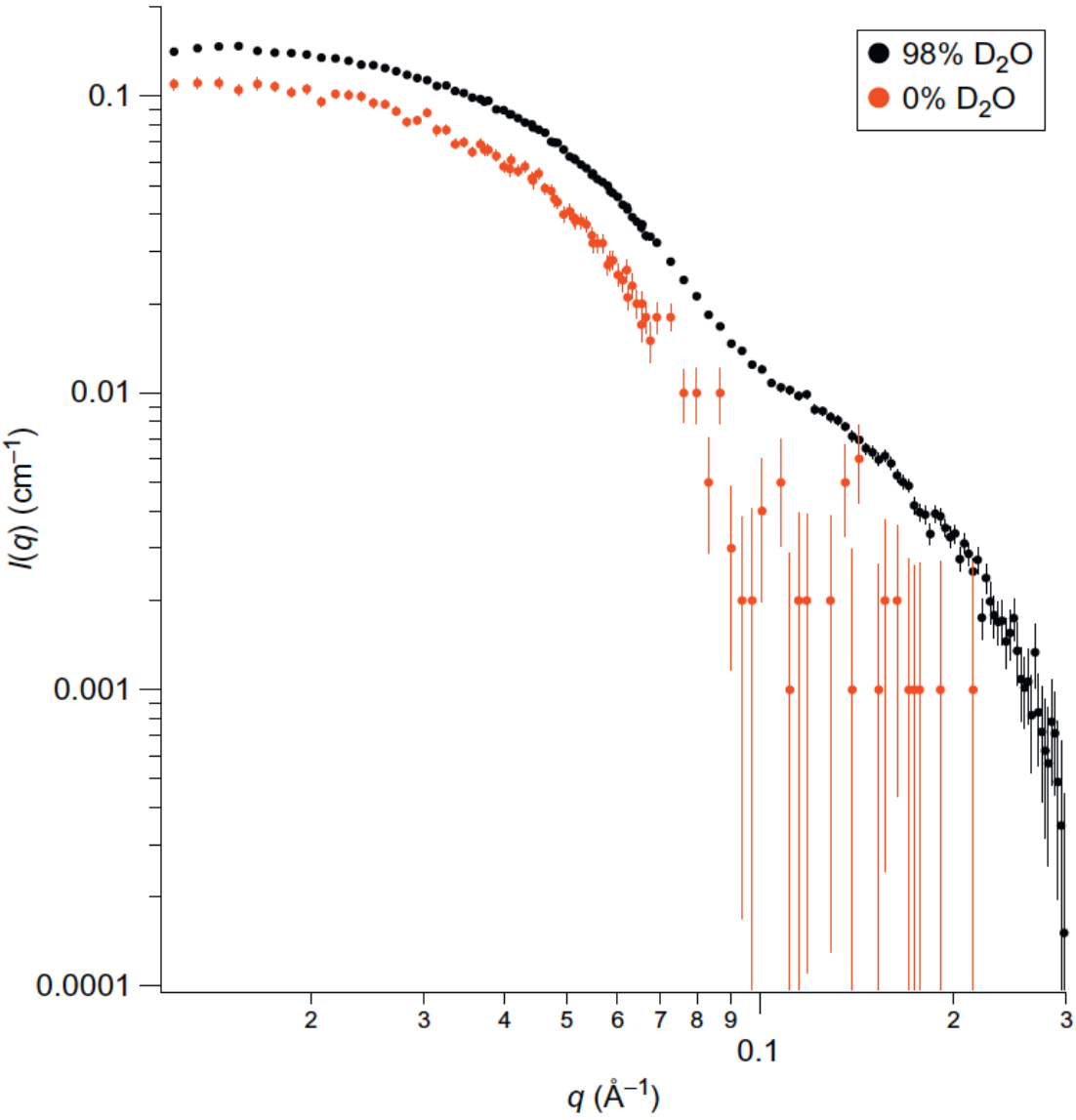


Figure 2.2. SANS data on an absolute scale for Skp in D_2O and H_2O buffers. Error bars represent the standard error of the mean with respect to the number of pixels used in the data averaging.

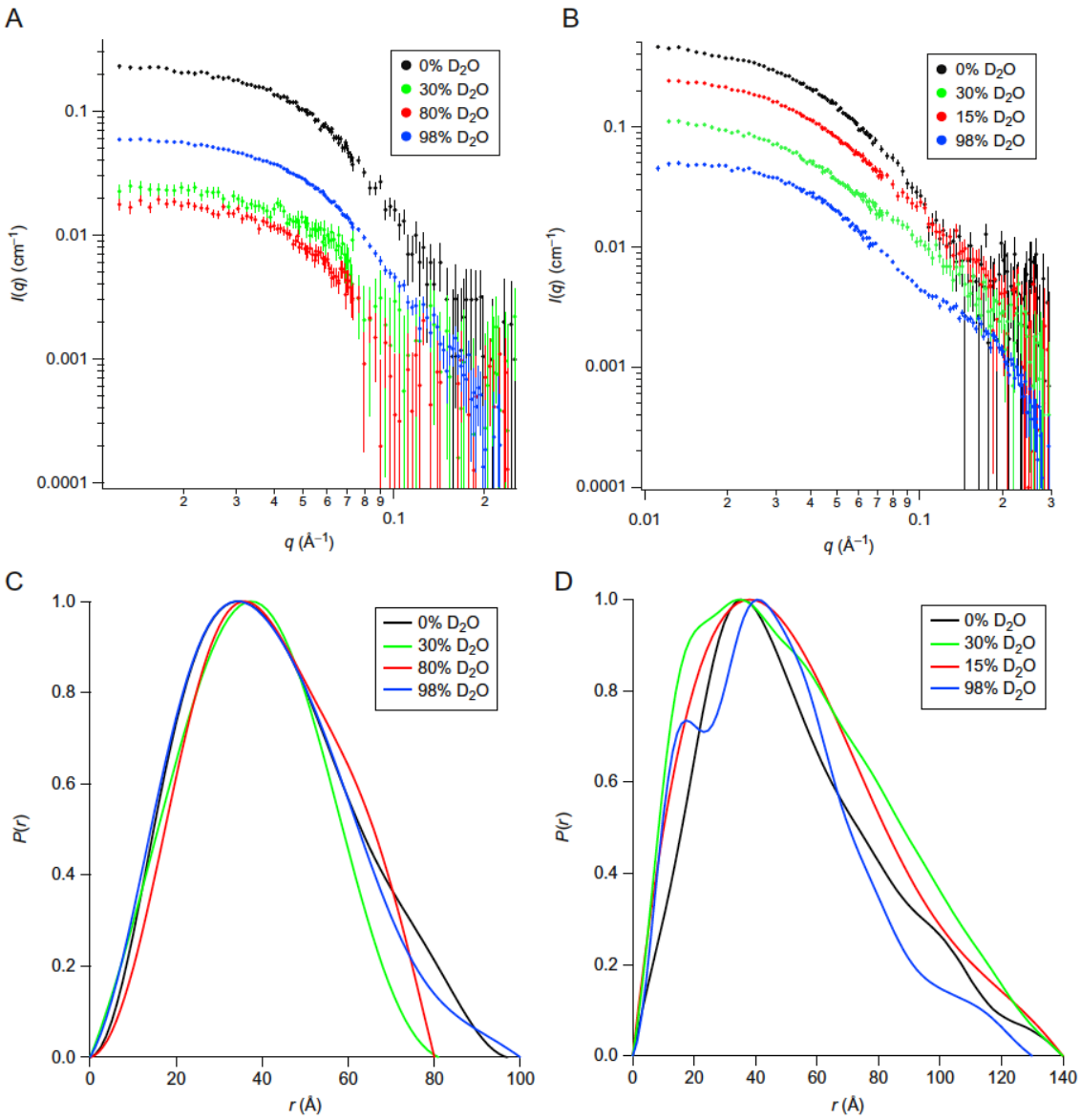


Figure 2.3. SANS data on an absolute scale for (A) Skp–OmpW and (B) Skp–OmpA. Error bars represent the standard error of the mean with respect to the number of pixels used in the data averaging. Distance distribution functions for (C) Skp–OmpW and (D) Skp–OmpA. The peak values are scaled to 1.0 so that the differences in shape can be easily observed.

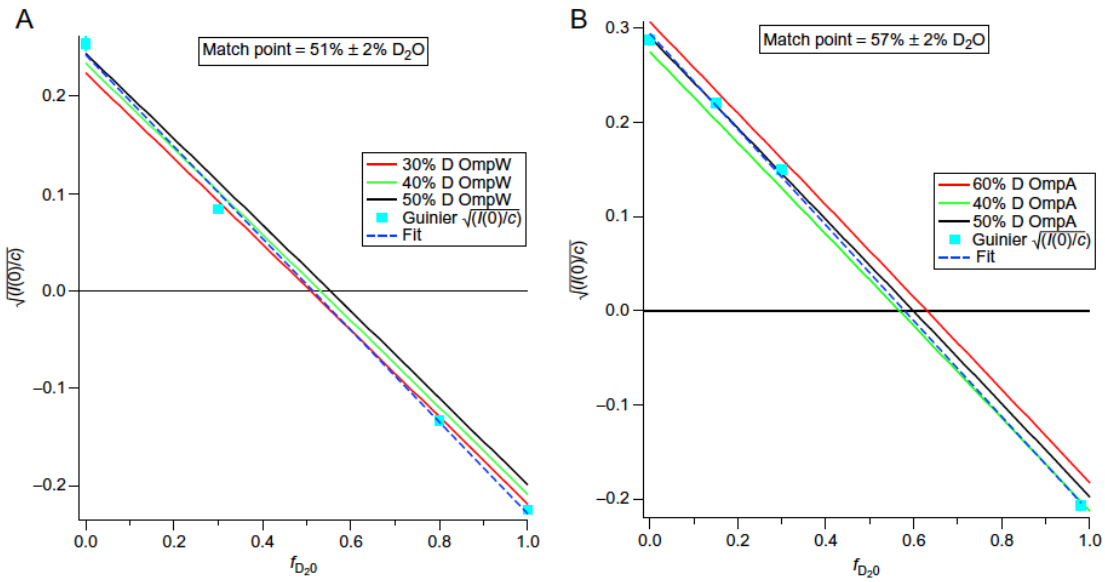


Figure 2.4. Experimental and calculated versus f_{D2O} for (A) Skp–OmpW and (B) Skp–OmpA. Error bars on the experimental values represent the standard error of the mean based on three different Guinier fits.

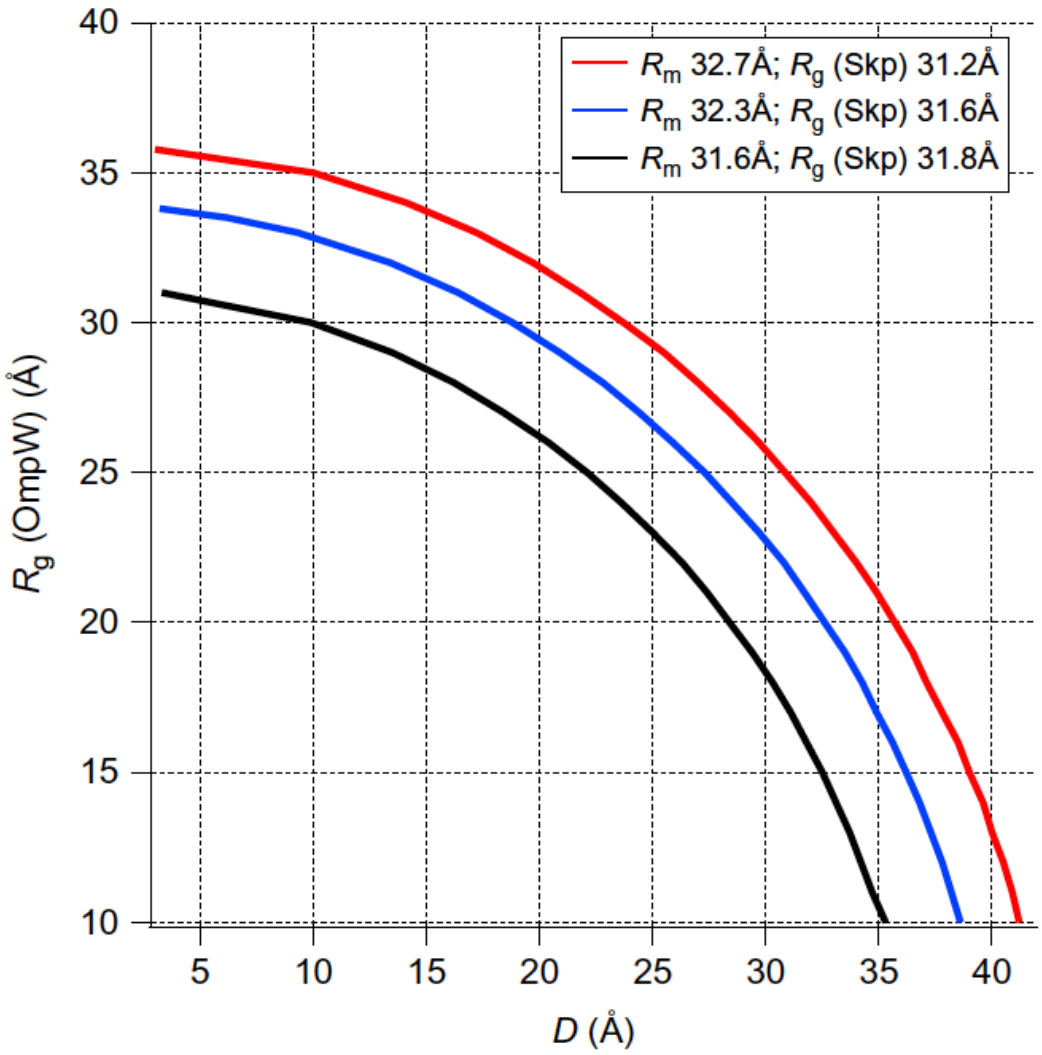


Figure 2.5. Calculation of possible values for R_g(OmpW) and CM distance, D_{CM}, based on R_m and R_g(Skp) obtained from the Stuhrmann analysis. Top curve: R_m = 32.7 Å; Middle curve: R_m = 32.3 Å; Bottom curve: R_m = 31.5 Å.

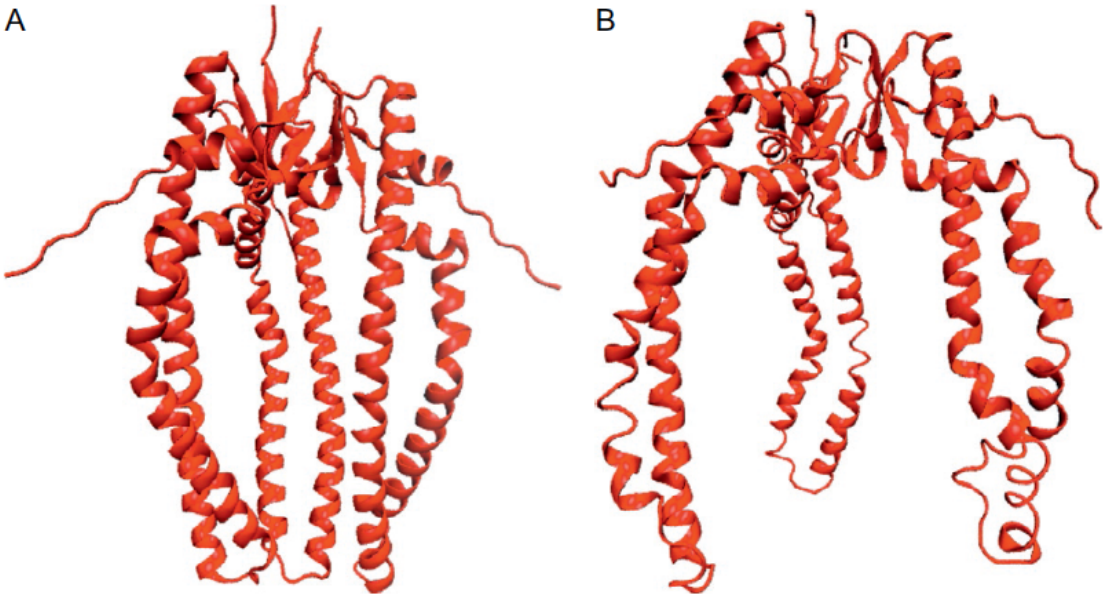


Figure 2.6. All-atom model structures representing Skp (A) in the crystal (R_g = 30 Å) and (B) in solution (R_g = 33 Å).

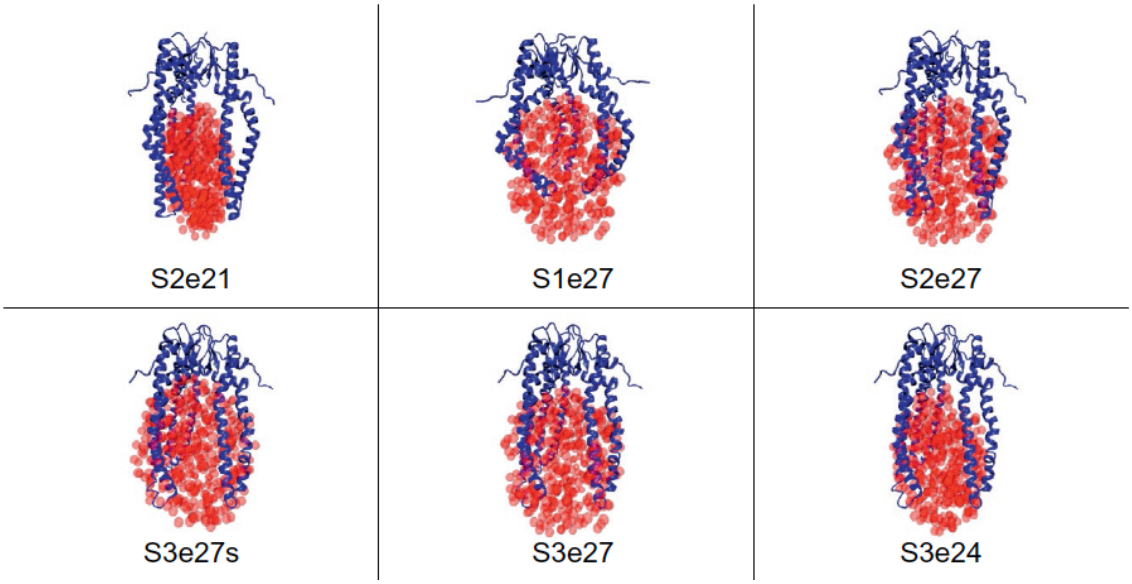


Figure 2.7. Skp–OmpW model structures tested against the SANS data as described in Section 2.3. The ball representation of the ellipsoids is for clarity. The ellipsoids were represented by nonoverlapping points for SANS curve calculations, as described in Section 2.3.

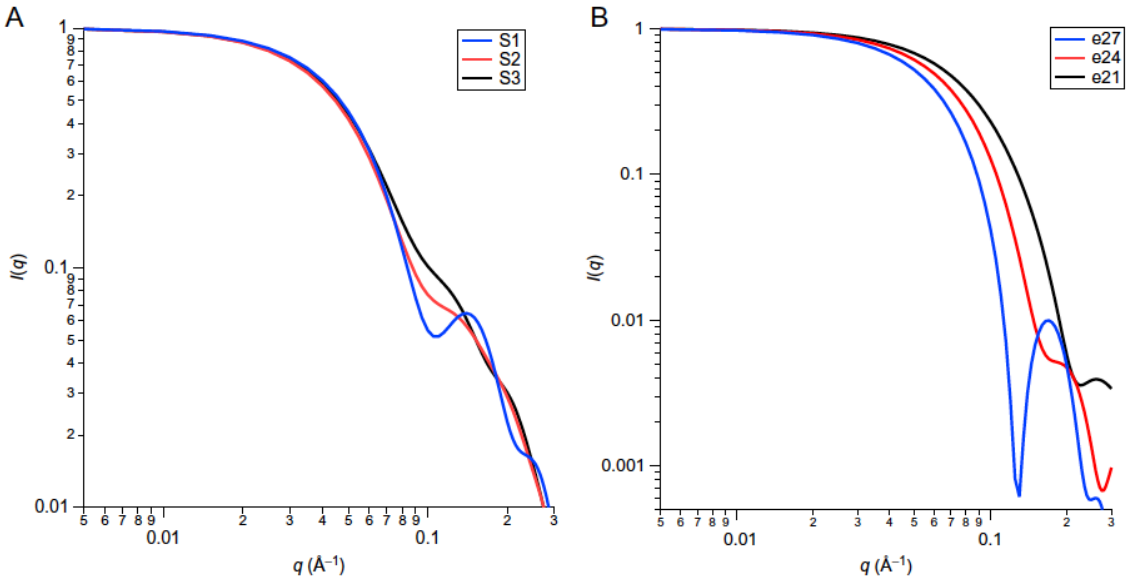


Figure 2.8. Model SANS curves from (A) the Skp component represented by the S1, S2, and S3 all-atom structures and (B) the OmpW component represented by the e21, e24, and e27 ellipsoids. I(0) is arbitrarily scaled to 1.0.

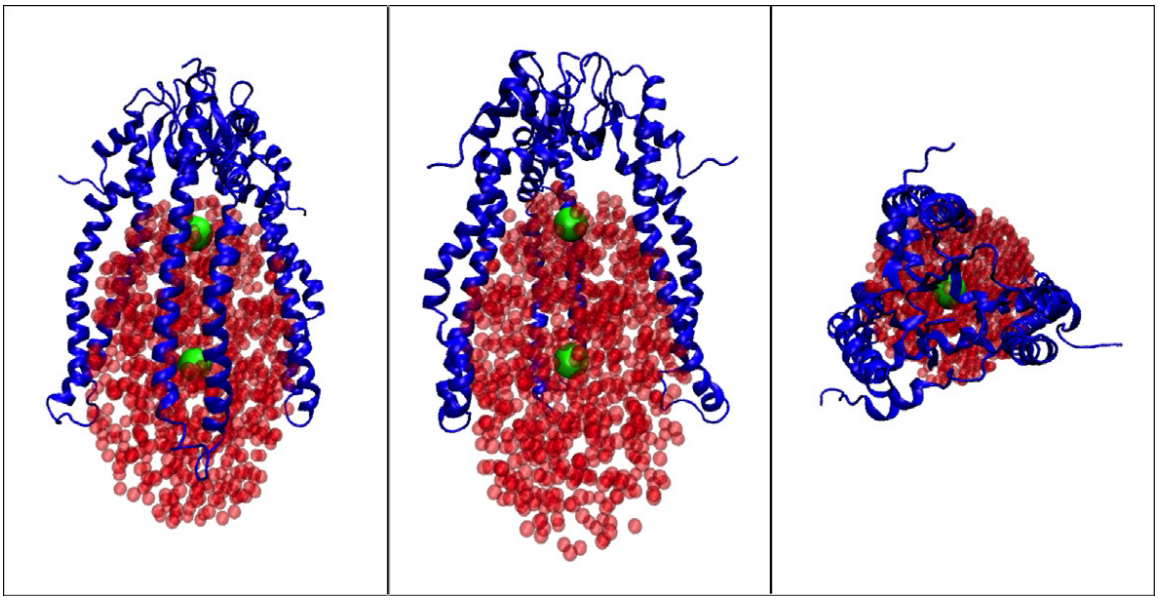


Figure 2.9. Several views of Skp–OmpW model S3e24, which was the best-fit model for this complex. The larger green (dark gray in the print version) balls are at the CM locations of S3 Skp and the e24 ellipsoid. The ball representation of the ellipsoids is for clarity. The ellipsoids were represented by nonoverlapping points for SANS curve calculations, as described in Section 2.3.

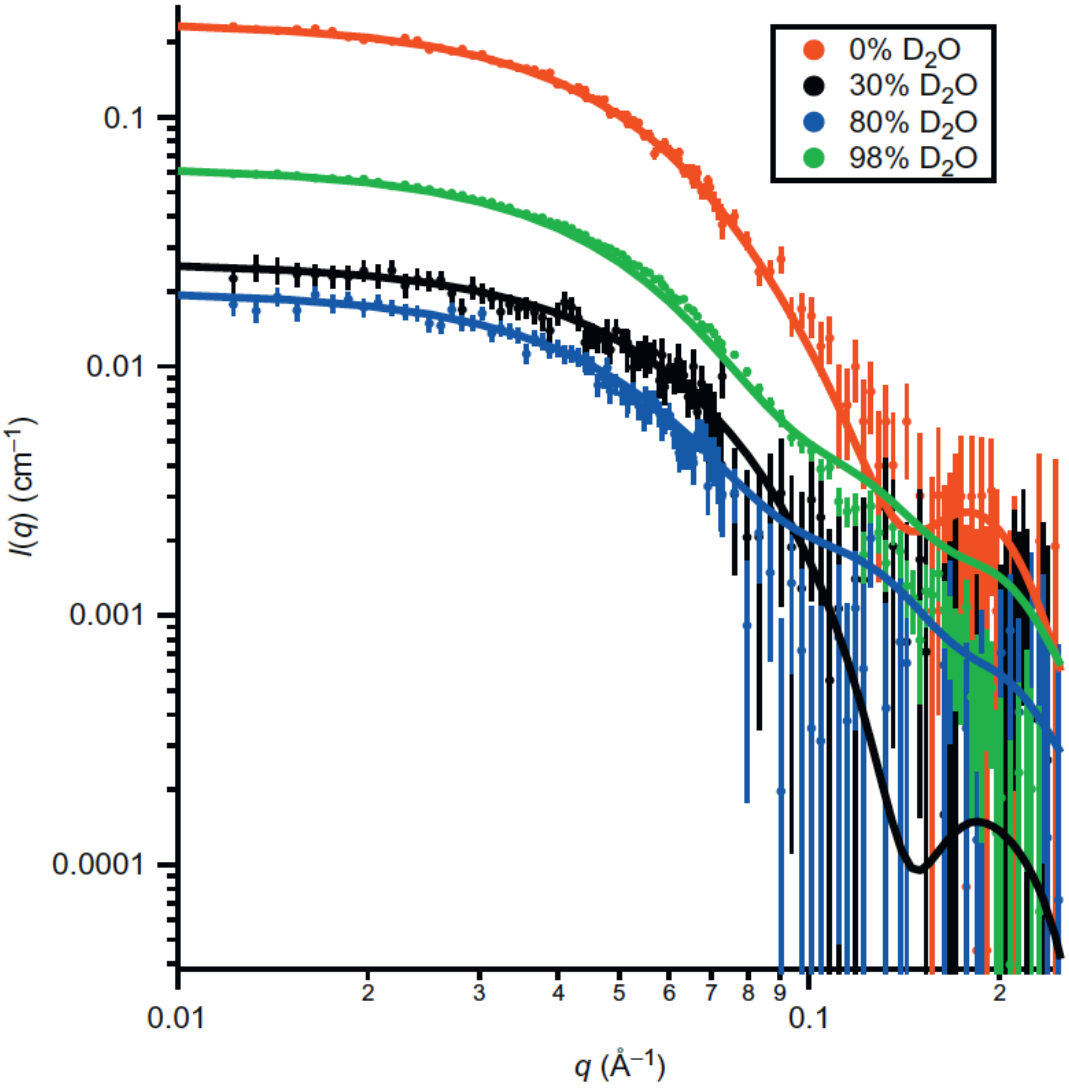


Figure 2.10. Skp–OmpW SANS data on an absolute scale (points) along with model SANS curves (solid lines) from S3e24. Error bars represent the standard error of the mean with respect to the number of pixels used in the data averaging.

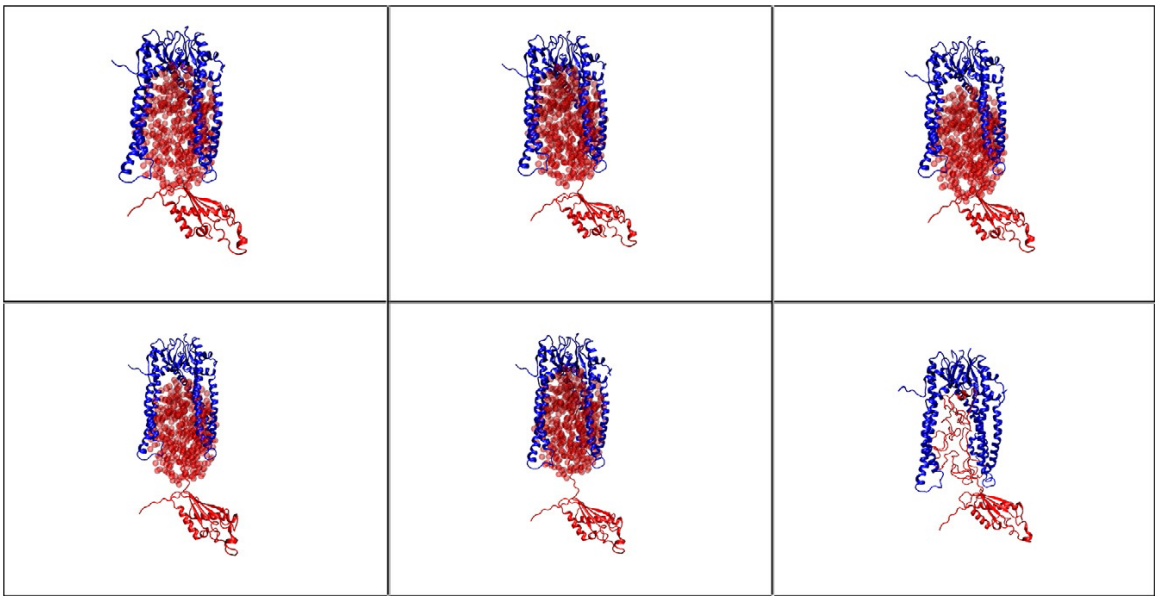


Figure 2.11. Skp–OmpA starting structure models for SASSIE runs as described in Section 2.3. The ball representation of the ellipsoids is for clarity. The ellipsoids were represented by nonoverlapping points for SANS curve calculations, as described in Section 2.3.

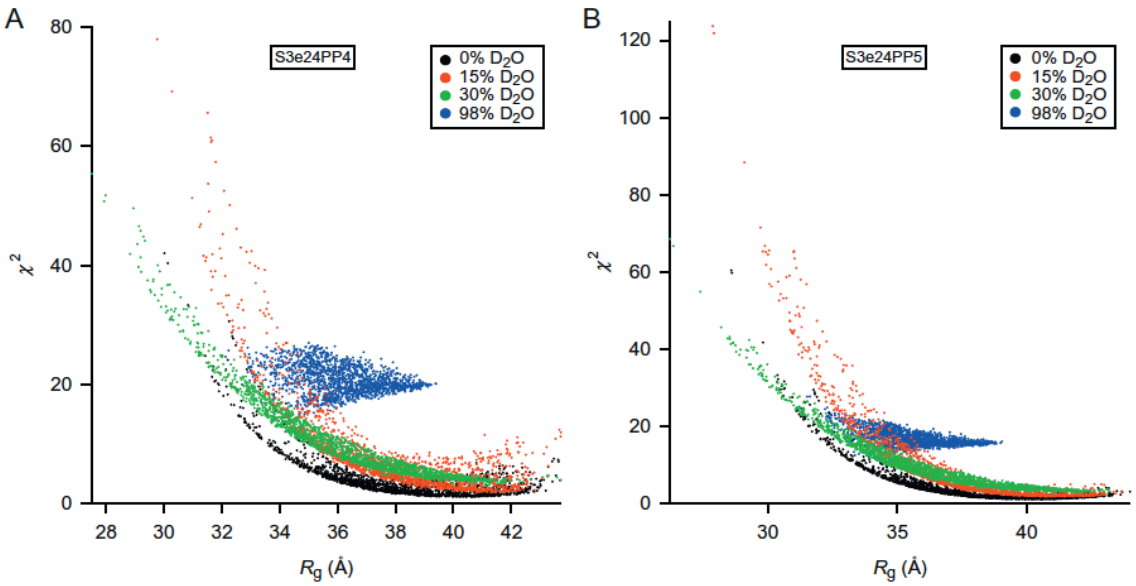


Figure 2.12. χ^2 versus R_g plots as a function of contrast from SASSIE runs exploring the conformation space of the OmpA PP domain for models (A) S3e24PP4 and (B) S3e24PP5.

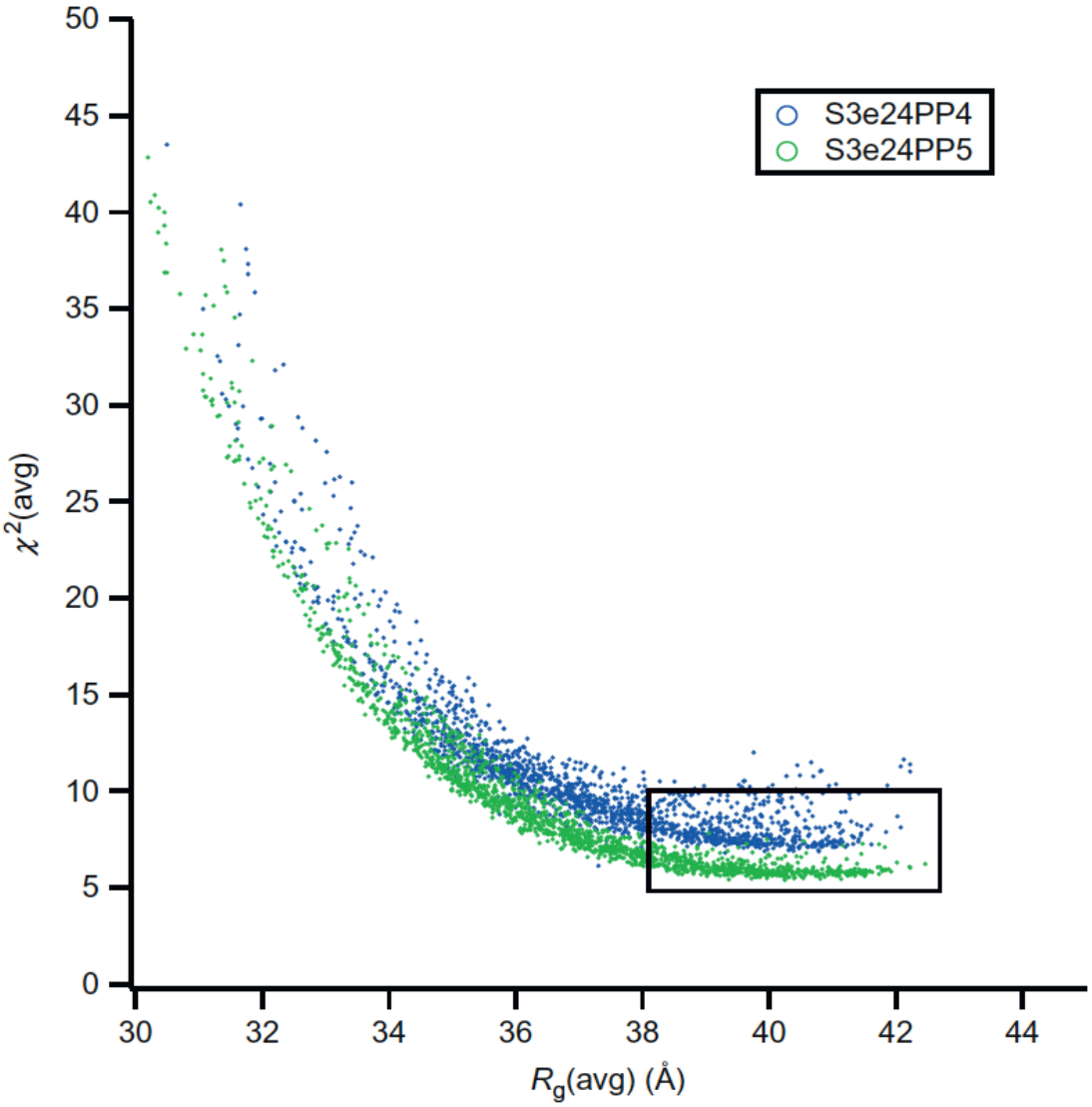


Figure 2.13. χ^2(avg) versus Rg(avg) plots from the plots in figure 2.12 for models S3e24PP4 (dark gray in print version) and S3e24PP5 (light gray in print version).

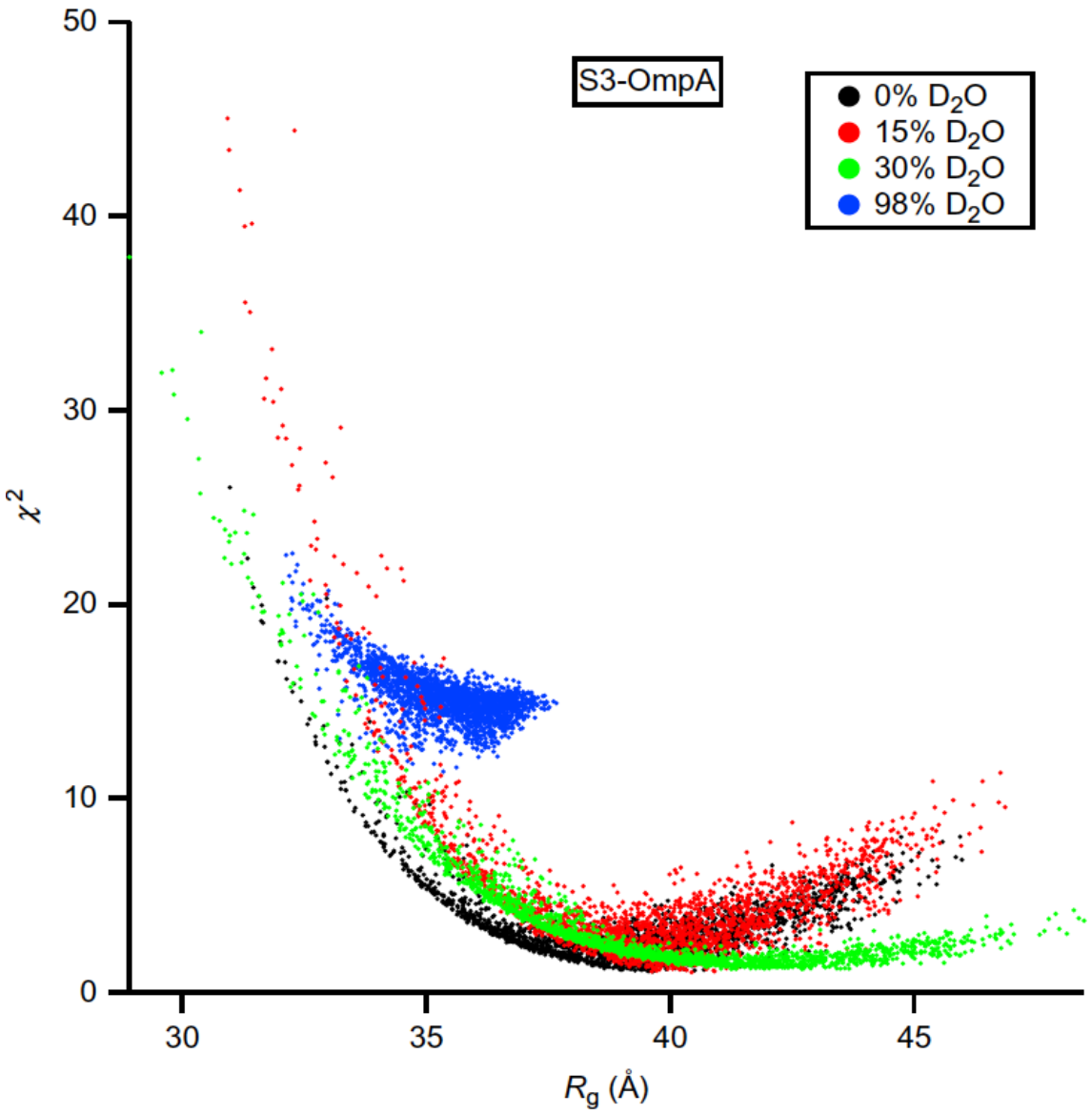


Figure 2.14. χ^2 versus R_g plots as a function of contrast from SASSIE runs exploring the conformation space of the OmpA PP domain for model S3-OmpA.

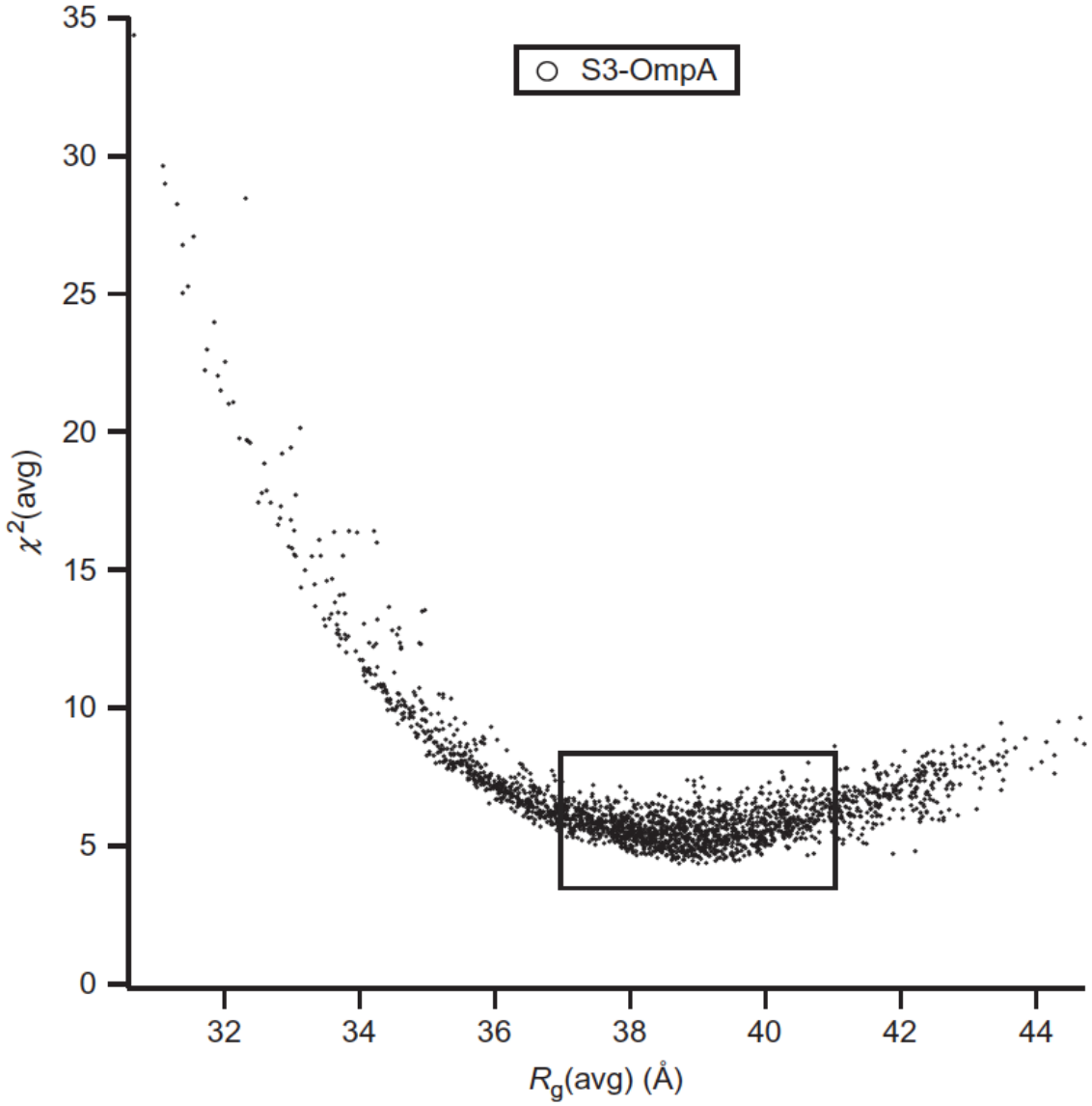


Figure 2.15. χ^2(avg) versus R_g(avg) plots from figure 2.14 for S3-OmpA.

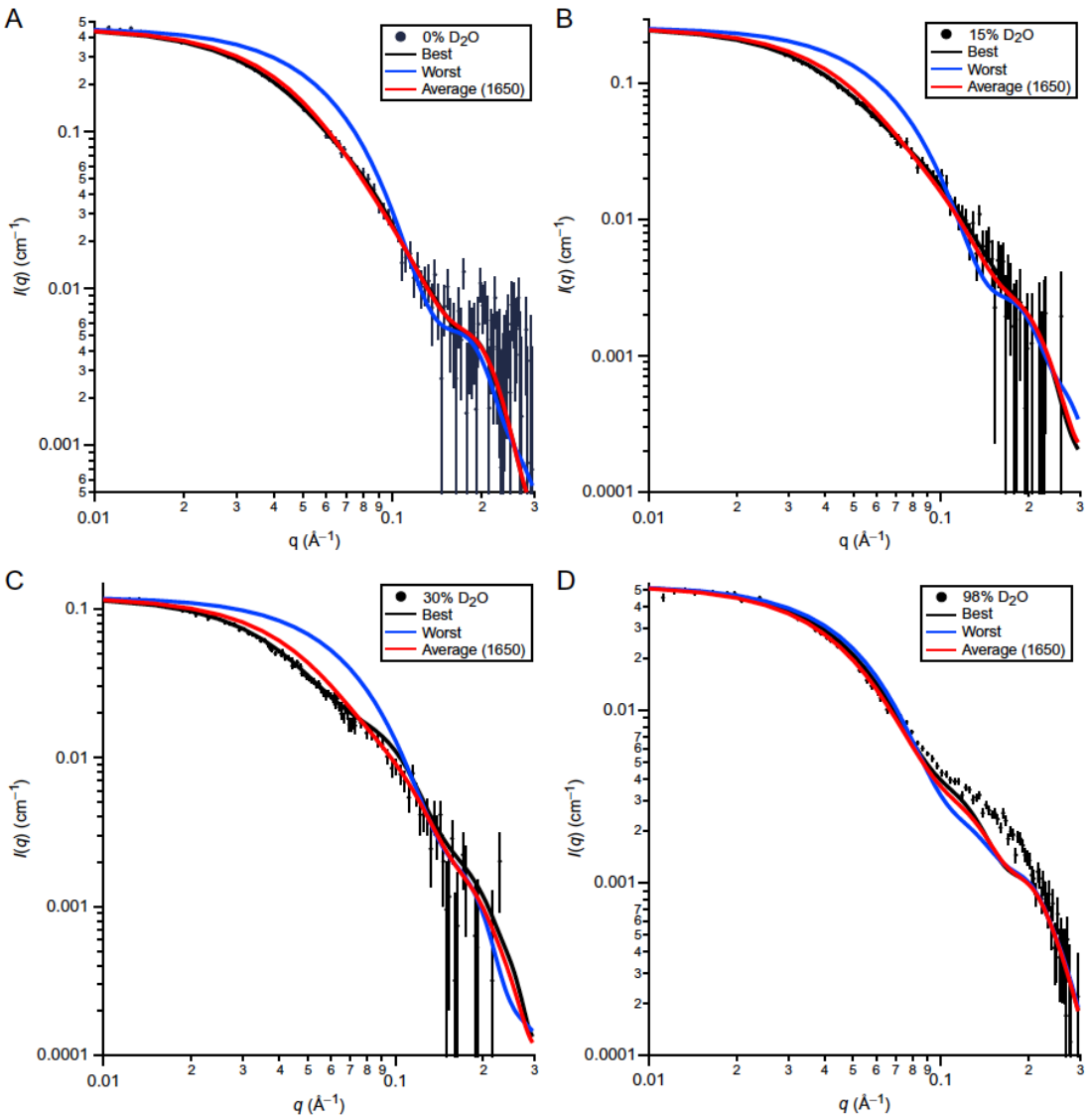


Figure 2.16. Model SANS data calculated from the best- and worst-fit single S3e24PP5 structures, as well as the average SANS curve calculated from the entire S3e24PP5 ensemble, along with the SANS data on an absolute scale for Skp–OmpA in (A) 0% D_2O, (B) 15% D_2O, (C) 30% D_2O, and (D) 98% D_2O. Error bars represent the standard error of the mean with respect to the number of pixels used in the data averaging.

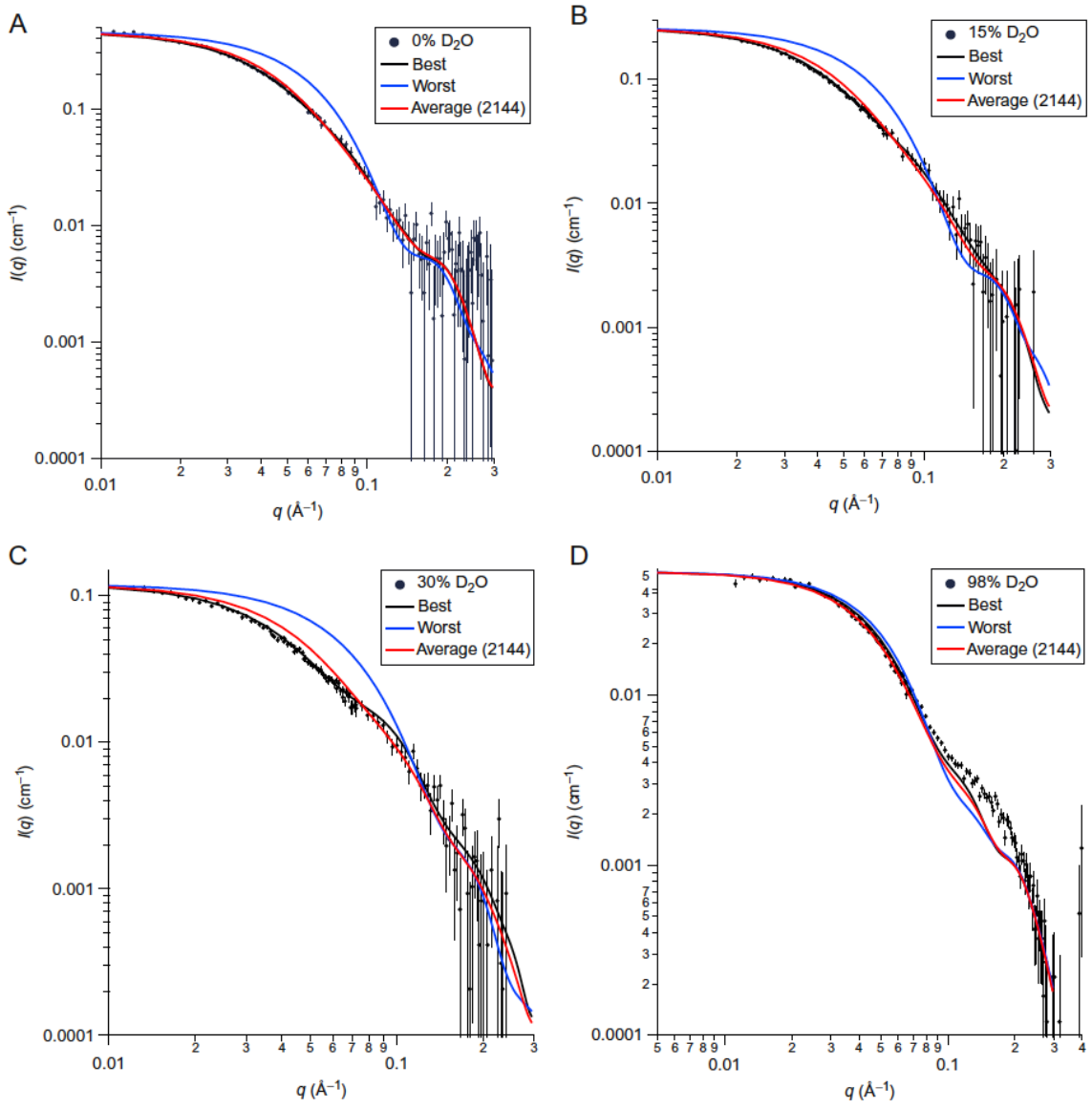


Figure 2.17. Model SANS data calculated from the best- and worst-fit single S3-OmpA structures, as well as the average SANS curve calculated from the entire S3-OmpA ensemble, along with the SANS data on an absolute scale for Skp–OmpA in (A) 0% D_2O, (B) 15% D_2O, (C) 30% D_2O, and (D) 98% D_2O. Error bars represent the standard error of the mean with respect to the number of pixels used in the data averaging.

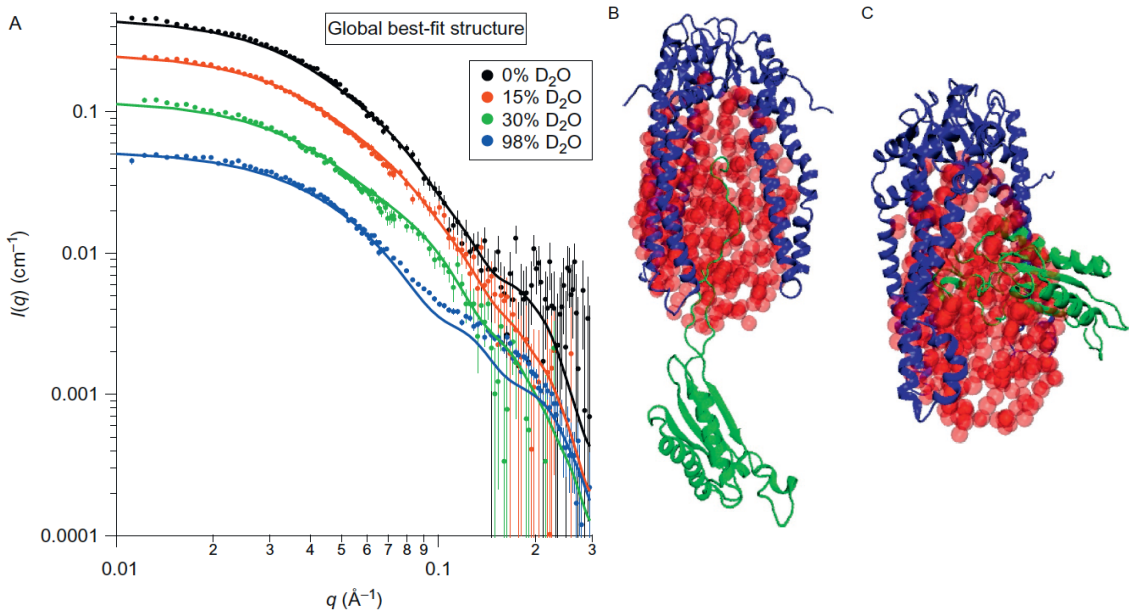


Figure 2.18. (A) Skp–OmpA SANS data on an absolute scale along with model SANS curves from the global single best-fit S3e24PP5 structure. Error bars represent the standard error of the mean with respect to the number of pixels used in the data averaging. The (B) best-fit and (C) worst-fit model structures are also shown for comparison. The ball representation of the ellipsoids is for clarity. The ellipsoids were represented by nonoverlapping points for SANS curve calculations, as described in Section 2.3.

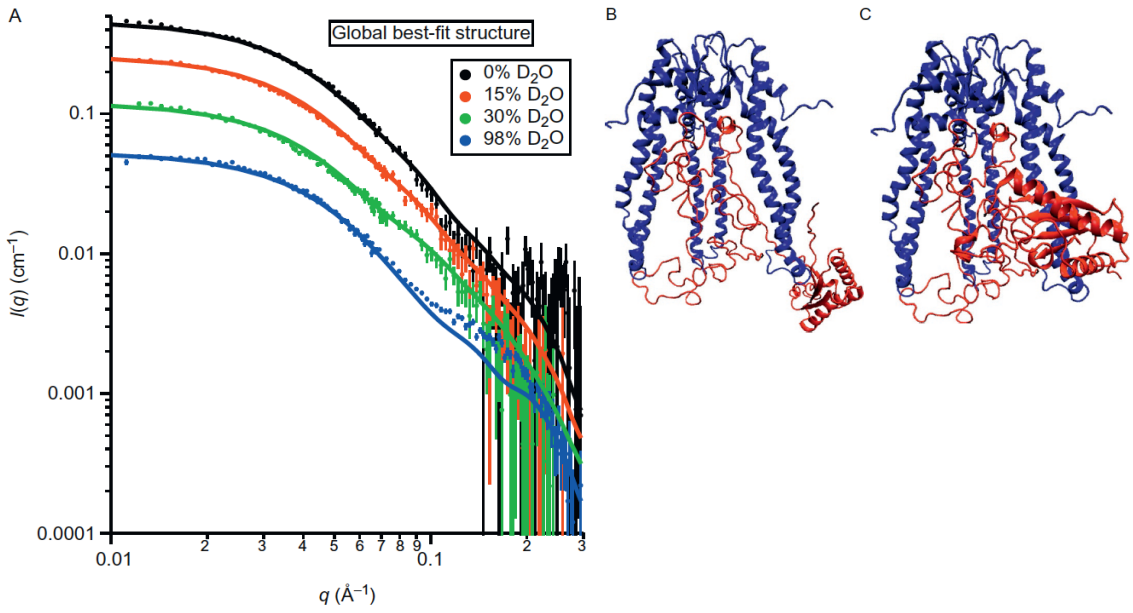


Figure 2.19. (A) Skp–OmpA SANS data on an absolute scale, along with model SANS curves from the global single best-fit S3-OmpA structure. Error bars represent the standard error of the mean with respect to the number of pixels used in the data averaging. The (B) best-fit and (C) worst-fit model structures are also shown for comparison. S3e24PP5 S3-OmpA

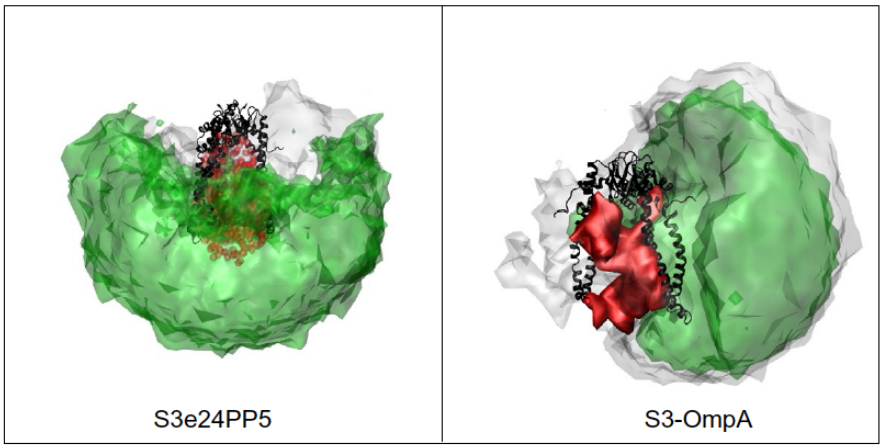


Figure 2.20. Structure density plots representing all of conformation space (gray (lightgray in print version)) and the best-fit conformation space within the rectangles in figures 2.13 and 2.15 (green (medium gray in the print version)) explored by the OmpA PPdomain using the S3e24PP5 and S3-OmpA models. The OmpA TM domain is shownin red (dark gray in the print version) and Skp is shown in black.

Chapter 3
Skp Chaperones OmpW, but Not OmpLA or BamA in Conditions which Favor Aggregation

3.1 Introduction

Transport of uOMPs across the periplasm from the translocon in the inner membrane (IM) to their points of assembly in the outer membrane (OM) requires chaperones. The protein Seventeen Kilodalton Protein (Skp) is thought to have chaperone activity (CA) with diverse uOMPs (see chapter 1). This idea is partly based on results that show Skp binds most uOMPs (Qu, Mayer et al., 2007, Moon, Zaccai et al., 2013) and inhibits their aggregation (Walton and Sousa, 2004, Walton, Sandoval et al., 2009, Entzminger, Chang et al., 2012). Furthermore, double deletion of Skp and the protease DegP results in increased formation of aggregates in the periplasm (Schäfer, Beck et al., 1999). These findings support the idea that Skp possesses an antiaggregation activity (AA) (figure 1.5), but they do not show that Skp is a chaperone for uOMPs.

Other groups have directly measured Skp CA, but their experimental methods vary (Bulieris, Behrens et al., 2003, Entzminger, Chang et al., 2012, McMorran, Bartlett et al., 2013, Patel and Kleinschmidt, 2013, Thoma, Burmann et al., 2015). Among these studies, Skp CA for uOMPs was only shown for OmpA, PagP and FhuA (Bulieris, Behrens et al., 2003, McMorran, Bartlett et al., 2013, Patel and Kleinschmidt, 2013, Thoma, Burmann et al., 2015). For FhuA, AFM showed that Skp did not assist folding (Thoma, Burmann et al., 2015). On the other hand, Skp CA for OmpA and PagP depended, variously, on LPS, BamA, or PG headgroups. Therefore, Skp CA may depend

on membrane components, or it may depend on the OMP. The latter idea, that CA depends on the substrate, is suggested by differences in the structures of Skp-uOMP complexes (chapter 2). Moreover, Skp CA may be limited for large OMPs that cannot fit within the Skp cavity (figure 1.7). Thus, I asked whether Skp CA varies for OmpW (8 β–strands), OmpLa (12 β–strands), and BamA (16 β–strands).

To test CA, I used SDS-PAGE (Nakamura and Mizushima, 1976) to quantify the folding efficiency of OMPs that were allowed to aggregate with or without Skp (Bulieris, Behrens et al., 2003). I show that in the presence of visible precipitation, a mostly (70%) soluble population of Skp dramatically improved the folding efficiency of uOmpW, but had the the opposite effect on uOmpLa, preventing its folding. For BamA, Skp had no effect. These results show that Skp may not possess CA for all uOMPs.

3.2 Materials and Methods

Delayed Folding Experiments

Figure 3.1 illustrates the protocol. To measure the CA of Skp, I subjected uOMPs to conditions promoting aggregation, called an "aggregation reaction," for time periods, called "delays," of 0 to 60 minutes. I performed these experiments with or without 22.1 μM Skp. Aggregation reactions consisted of 4.9 μM uOMP, 20 mM Tris-HCl, 19.3 mM NaCl and 1 M urea (Sigma), pH 8.0. After the delay, the aggregation reaction was combined with large unilamellar vesicles (LUVs) to start a folding reaction. LUVs were 1,2-didecanoyl-sn-glycero-3-phosphocholine ($diC_{10}PC$) and 1,2-didecanoyl-sn-glycero-3-phospho-(1'-rac-glycerol) ($diC_{10}PG$) in a 4:1 molar ratio of PC:PG (Avanti Polar Lipids). LUV solutions contained 20 mM Tris-HCl, 20 mM NaCl and 1 M urea at pH 8.0. The

final folding reaction conditions were 4 μM uOMP, with or without 18 μM Skp (1:4.5 uOMP:Skp), in 20 mM Tris-HCl, 19.4 mM NaCl, 1 M urea, and 3.2 mM LUVs (800:1 lipid:uOMP ratio), pH 8.0. Extrusion of LUVs was performed as described (Burgess, Dao et al., 2008), with 11 passes through the filter.

Folding reactions were incubated at ambient temperature for 2 hours with stirring. Reactions were quenched with a modified 4X Laemmli loading buffer containing 32% glycerol and 40 mM β-mercaptoethanol. After taking the last sample, each sample was divided into two aliquots, and one of the two was heated at 95 °C for 5 minutes. 14 μL of each sample was loaded in a precast Bio-Rad Mini-Protean gel with 12% polyacrylamide concentration. SDS-PAGE was performed in a Bio-Rad Mini-Protean system and gels were scanned using a desktop scanner. Densitometry was performed with ImageJ software.

Sedimentation Velocity of Skp to Assess Solubility as a Function of Urea Concentration

Skp at 4 °C and less than 150 μM concentration was incubated at room temperature for 15 minutes. This Skp was first diluted to a concentration of 60 μM in gel filtration buffer (GF, 20 mM Tris-HCl and 200 mM NaCl at pH 8.0). This predilution was further diluted with a mixture of urea stock (20 mM Tris-HCl, 6.0 M urea, pH 8.0) and aqueous buffer (20 mM Tris-HCl, pH 8.0). Final concentrations were 40 μM Skp, urea concentrations of 80, 200, 300, or 500 mM, and 150 mM NaCl. The first data point for Skp solubility in figure 3.5, at 0 mM urea corresponds to Skp at 40 μM concentration in GF. I assume NaCl does not affect Skp solubility, because it was shown (Sandlin,

Zaccai et al., 2015) that Skp solubility in SE is insensitive to NaCl at ambient temperature for several days.

The reaction was allowed to proceed without stirring while being loaded at room temperature into a 2-sector centerpiece. After incubating for 30 minutes at 25 °C, the centrifuge was started. All SV experiments were performed using a Beckman XL-A AUC with an An60-Ti rotor, at 201,600 g. The step-size for data collection was 0.005 cm, with data collected at 235 nm. The late-sedimenting species was identified using DCDT+ software by Philo (2006), and this data was fit to a single Gaussian (figure 5.2A). The absorbance contribution for the monomeric species was estimated from the fitted value C0 for the monomer. To determine the total absorbance, the average of three values was taken near 6.5 cm from the first scan.

3.3 Results

Delayed-Folding Into LUVs Occurs in the Presence of OMP Precipitates

In all aggregation reactions, I observed varying amounts of turbidity consistent with formation of aggregates (Danoff and Fleming, 2015). Because it was previously shown that aggregation reduces folding efficiency for OmpA in delayed folding assays (Danoff and Fleming, 2011), the degree of turbidity was not further quantified. All results are thus interpreted as relative measures of CA in the context of a nonhomogeneous mixture of folded, unfolded, and aggregated components. For all other experiments I have performed with Skp and are described in chapters 2, 4, and 5, no visible precipitation was observed.

OmpW and OmpLa Folding Efficiency Decreases with Increasing Delay in the Absence of LUVs

Lipids are known to be required for OMP folding to occur. Previously, it was shown (Danoff and Fleming, 2011) that as the delay time increases before LUVs are added, the folding efficiency of uOmpA decreases. Using delayed-folding (section 3.2 and figure 3.1), I tested whether the same loss of folding efficiency occurred for OmpW, OmpLa, and BamA. After allowing aggregation reactions to proceed for from 30 seconds to 1 hour, mixtures were added to diC_{10} LUVs with 80% PC/20% PG headgroups at a final ratio of 800:1 lipid:protein. After allowing OMPs to fold for two hours in the LUVs with 1 M urea and pH 8.0, these reactions were quenched (section 3.2).

As shown in figures 3.1 and 3.2, folding efficiency decreases gradually for OmpW and OmpLa as the delay time without lipids increases. This contrasts with BamA, for which folding efficiency was poor and independent of delay. This pattern might be explained by a rapid loss of folding efficiency over the initial 30 seconds. To test this possibility, I added BamA directly to LUVs (0 time point, figure 3.2), but saw the same poor folding efficiency. Therefore, BamA aggregation may occur rapidly and irreversibly during mixing, with or without the presence of LUVs. With this assumption, the kinetics of decay of OMP folding efficiency appears to follow the trend uBamA > uOmpW > uOmpLa.

Skp Possesses Chaperone Activity with uOmpW, but not for uOmpLa or uBamA

To test whether Skp CA differs between uOMPs, I performed delayed-folding with Skp present at a 4.5:1 Skp:uOMP ratio in the aggregation reactions. As shown in

figures 3.2 and 3.3A, after transferring the aggregation reactions to lipids, folding yields for OmpW increased in the presence of Skp. This increase occurred for delay times of up to 20 minutes. Interestingly, Skp had the opposite effect for OmpLa, decreasing the folding yield. No effect of Skp was observed for BamA folding.

I next asked whether Skp prevented the decay in folding efficiency over time, as expected for a holding chaperone. To answer this question, I examined the normalized folding efficiency, defined as the absolute folding efficiency divided by the maximum folding efficiency. For uOmpLa and uBamA, the normalized data show that Skp had no significant effect on the relative decay in folding efficiency (figures 3.2E and 3.2F). Taken together, I conclude that Skp has holding CA with uOmpW, but not with uOmpLa or uBamA at these conditions and time scales.

Because Skp unfolds in 1 M urea (McMorran, Bartlett et al., 2013), I thought it important to rule out that Skp aggregation accelerates OmpLa aggregation. If Skp promotes uOmpLa aggregation, there should be a different rate of decay in the normalized data for uOMPLa with and without Skp. However, the normalized data for OmpLa show the same pattern with or without Skp (figure 3.3E). Thus, Skp does not promote aggregation of uOmpLa, but it does prevent 80% of the OmpLa from folding (figure 3.2B). Because Skp binds uOmpLa tightly (Moon, Zaccai et al., 2013), the remaining 20% of OmpLa that folds (figure 3.2B) could be free OmpLa that is not bound to Skp.

70% of Skp is Functionally Folded in 1 M Urea

With the observation that "leaky" folding of free uOmpLa may be occurring, I wished to quantify the fraction of Skp which is active in these experiments at 1 M urea. Skp is known to unfold in 1 M urea and is estimated by McMorran et al. to have a sharp unfolding transition midpoint of 1.4 M urea (McMorran, Bartlett et al., 2013). To quantify folded Skp as a function of urea, I used SV to quantify the solubility of Skp as a function of urea concentration (section 3.2). Figure 3.5 shows the dependence of folded Skp on urea, $f_{\text{trimeric Skp}} = 0.96 \pm 0.015 - (0.23 \pm 0.053)[\text{Urea, M}]$. Extrapolating this value to 1 M urea predicts that up to 70%, of Skp is functionally folded in these delayed-folding assays.

3.4 Discussion

Many studies of Skp refer to it as a chaperone (chapter 1), but there are only a few studies that address whether Skp has CA for OMPs. Entzminger et al. tested Skp CA with three antibody fragments, and found that Skp inhibited folding for one, had little effect on the second, and improved folding for the third (Entzminger, Chang et al., 2012). Here, with three natural Skp substrates (Chen and Henning, 1996), I observed exactly the same pattern of CA response.

The identical pattern of CA response in two independent studies suggests Skp CA is highly sensitive to substrate type (Bulieris, Behrens et al., 2003, Patel and Kleinschmidt, 2013, Thoma, Burmann et al., 2015). Therefore, although the evidence is clear that Skp binds many disordered protein sequences (Entzminger, Chang et al., 2012, De, Jeong et al., 2014) and has AA (Schäfer, Beck et al., 1999, Wu, Ge et al., 2011,

Entzminger, Chang et al., 2012, Lyu, Shao et al., 2012), Skp only displays CA with some substrates (figure 3.4).

Those substrates do not include OmpLa. Skp prevented 80% of OmpLa from folding in the presence of PG headgroups. This is surprising, because PG headgroups activated Skp CA for both OmpA and PagP (Patel, Behrens-Kneip et al., 2009, McMorran, Bartlett et al., 2013). Therefore, activation of Skp CA by PG could be limited to small OMPs like OmpA, PagP, and OmpW. Alternatively, the lack of OmpLa folding could be influenced by slow release of uOmpLa (Burmann, Wang et al., 2013, Thoma, Burmann et al., 2015) to the membrane (figure 1.5).

The Skp cavity is small (Walton, Sandoval et al., 2009), with just enough space to accommodate the 8-stranded OMPs (figure 1.7). Of the three uOMPs I assayed, OmpW is small enough to fit in this cavity (Burmann, Wang et al., 2013). Also, Skp showed CA with uOmpW for up to 20 minutes. For the uOMP of intermediate size, uOmpLa, the results are best explained by tight binding without CA, whereas for the largest OMP, BamA, the data show no effect. This trend with increasing OMP size suggests that future Skp CA experiments should utilize ratios of (Skp:uOMP) mass concentration to complement molar concentration.

3.5 Figures

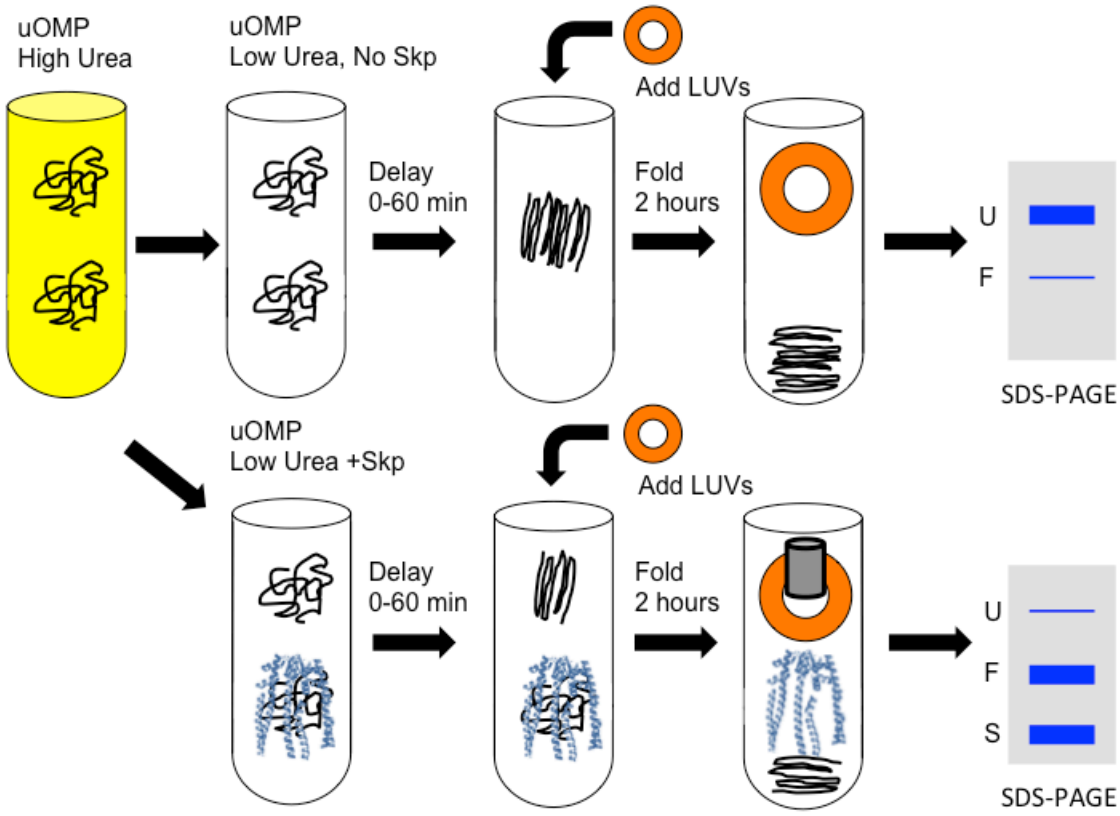


Figure 3.1. Schematic Representation of the Delayed-Folding Protocol (section 3.2). Top row, rapid dilution of OMPs from high urea (yellow), leads to aggregation. After addition of LUVs, folding reactions proceed for 2 hours and are then assayed by SDS-PAGE. U, unfolded band, F, folded band. Bottom row, the same experiment in the presence of Skp. If Skp is a chaperone, it should prevent aggregation and promote folding (compare the size of F in the bottom row to F in the top row). S, Skp also appears on the gel.

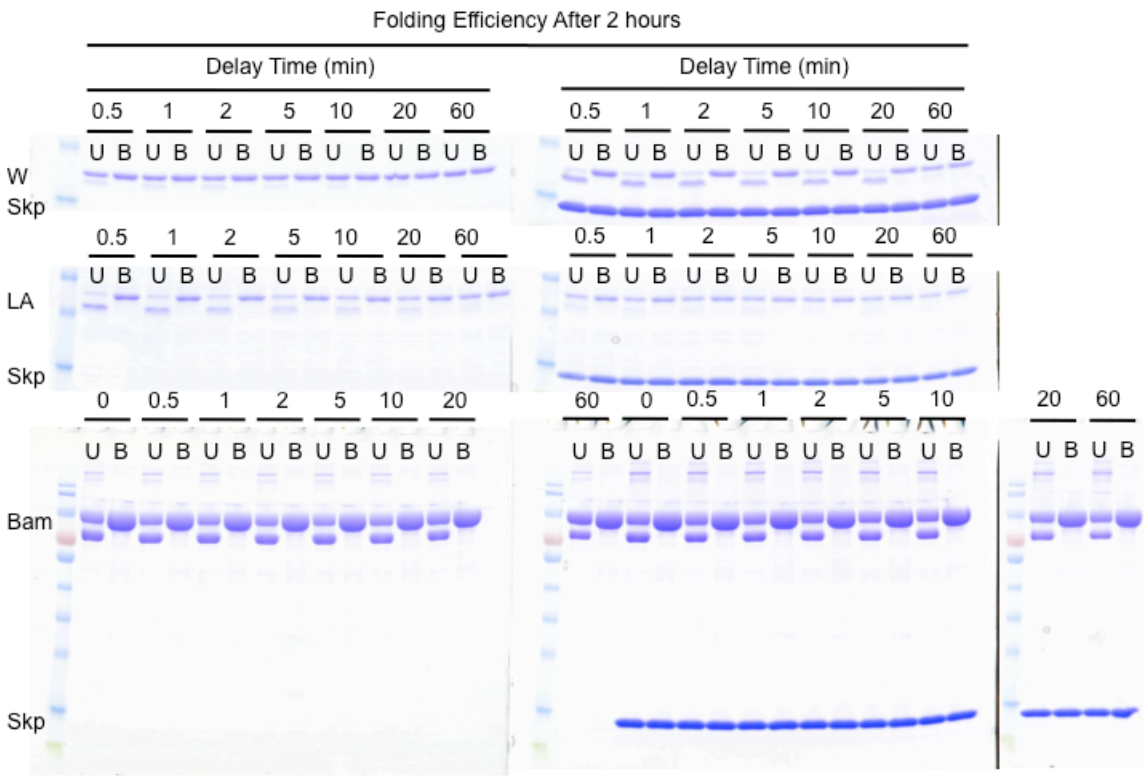


Figure 3.2. Delayed-Folding SDS-PAGE Gels Show Skp Has Chaperone Activity for OmpW, but Not OmpLa or BamA in the Presence of Aggregation. After rapid dilution of uOMPs without lipids (pH 8.0 and 1 M urea) and a delay period of from 0.5 to 60 minutes, uOMPs were added to LUVs and incubated for 2 hours. U, not boiled, B, boiled. Top, OmpW (W in the figure) delayed folding without Skp (left) shows very little folding even at short times, but with Skp present at a 4.5:1 molar ratio, right, folding is greatly improved. Middle, OmpLa (LA in the figure) 2-hour folding efficiency is inhibited by Skp. Bottom, BamA (Bam in the figure) folding is independent of delay and Skp. To show that no significant change occurs in the initial 30-second delay for BamA, a control was performed in which BamA was added directly to lipids (Delay Time 0), either alone or with Skp.

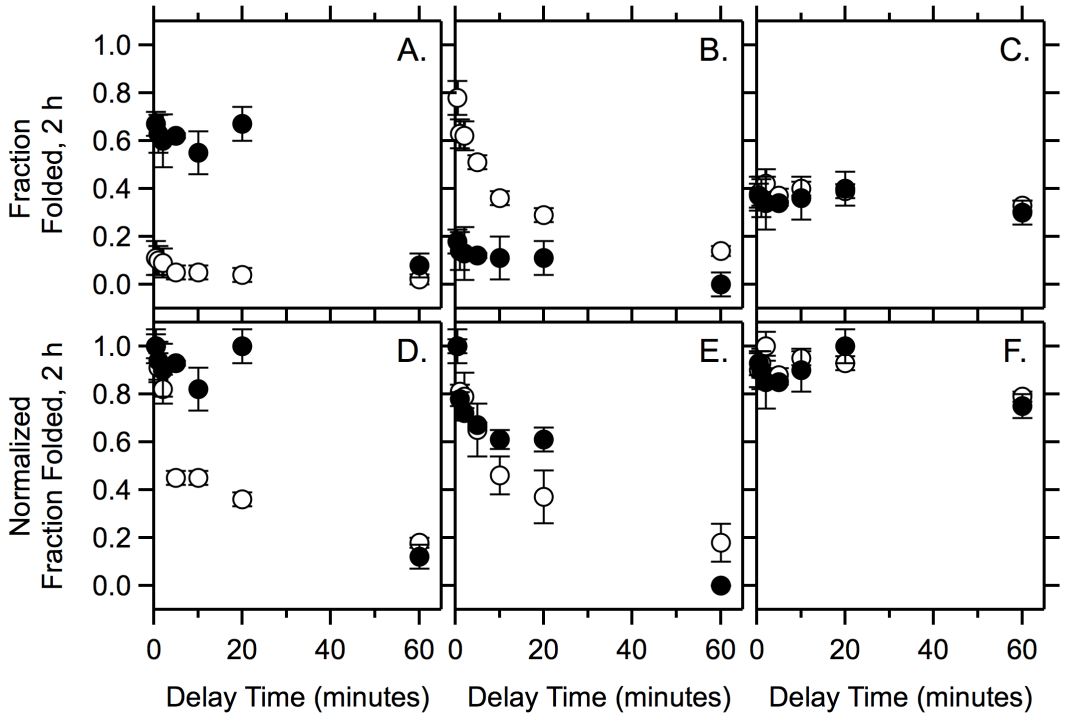


Figure 3.3. Quantification of Delayed Folding Data for uOmpW, uOmpLa, and BamA With and Without Skp. Filled circles, with Skp, open circles, without Skp. A., Skp shows chaperone activity with uOmpW for up to 20 minutes. B., uOmpLa loses absolute folding effiency owing to aggregation, whether Skp is present or absent (compare to A.). With Skp present, OmpLa folding efficiency is reduced ~80% at all delay times. C., BamA folding efficiency is poor even if added directly to lipids, and is unaffected by Skp. D., data in A-C., normalized to emphasize changes over time. E, normalized data from B show that Skp does not accelerate aggregation for uOmpLa.

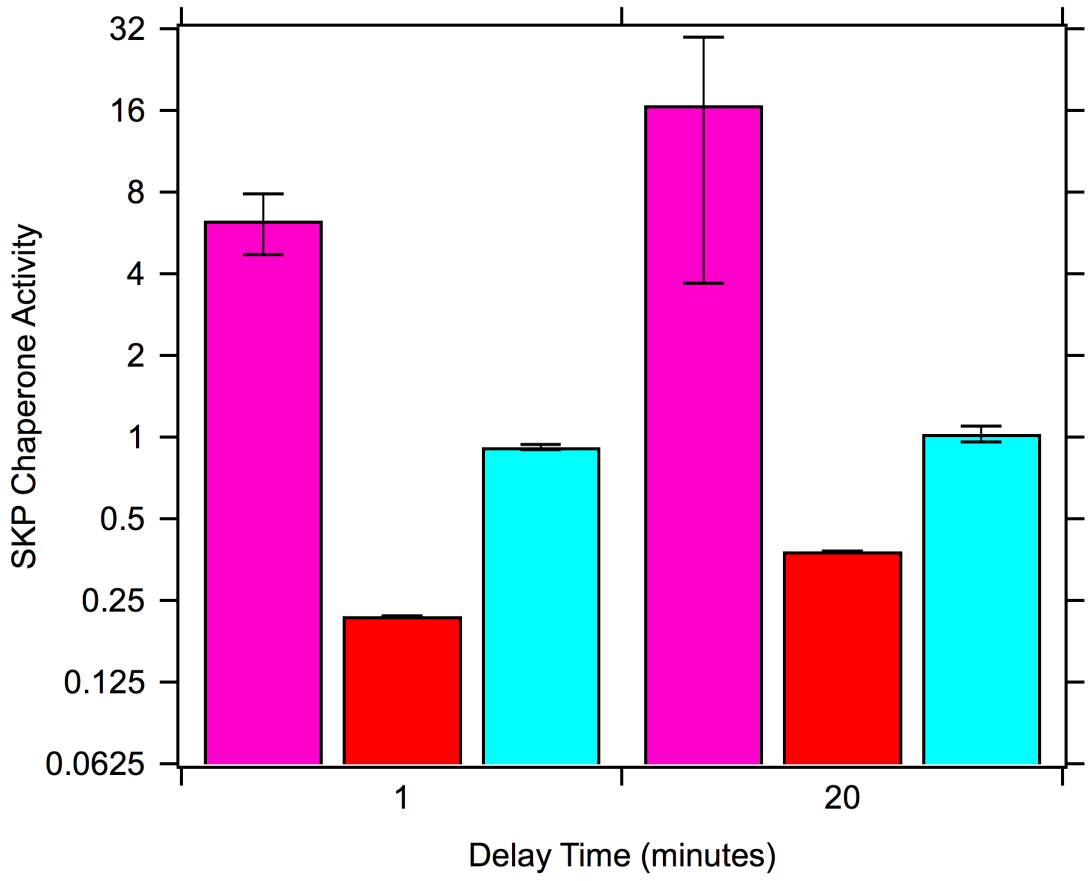


Figure 3.4. Skp Chaperone Activity Expressed as Fold-change in Folding Efficiency. Magenta, at both one and 20 minute delays without lipids, Skp dramatically improves 2-hour folding efficiency of OmpW, magenta, into 20% PG lipids. Red, Skp inhibits OmpLa folding efficiency. Cyan, Skp has no effect on BamA folding efficiency.

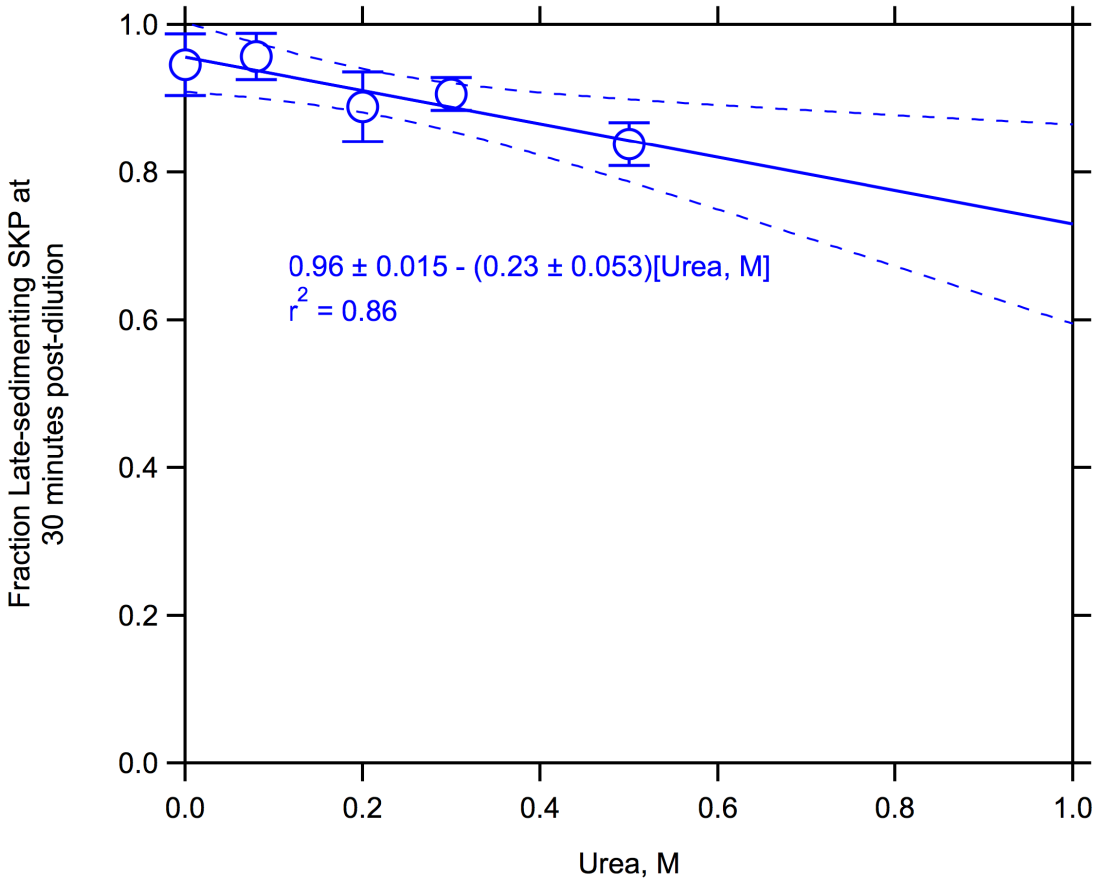


Figure 3.5. Skp Solubility as a Function of Urea Concentration. The Skp soluble fraction in NaCl (200 mM for 0 mM urea, 150 mM for all other urea concentrations) and, left to right, 0, 80, 200, 300, and 500 mM urea at pH 8.0. Data were collected by SV at 25 °C and 201,600 g, 30 minutes after dilution of stock into urea. Regression line, $f_{trimeric\ Skp} = 0.96 \pm 0.015 - (0.23 \pm 0.053)[Urea]$. Upper and lower lines represent 95% confidence bands for the linear regression.

INTENDED TO BE BLANK

Chapter 4

Skp Trimer Formation is Insensitive to Salts in the Physiological Range

Clifford W. Sandlin, Nathan R. Zaccai, and Karen G. Fleming*

T. C. Jenkins Department of Biophysics, The Johns Hopkins University, Baltimore, Maryland 21218, United States

*Corresponding author, email karen.fleming@jhu.edu

Preface: Changes to the Published Manuscript

This chapter is included as it was published in the journal Biochemistry, volume 54, 2015, except for the following changes. All sections were separated into sections to remain consistent with the format of the rest of the thesis. Also, in response to a comment from the committee, I have altered the phrase "…a negligible role for electrostatics in Skp trimerization." to "…a negligible role for long-range electrostatics in Skp trimerization." Finally, the chemical reaction notation is changed. The subscripts S_m and S_t for Skp monomers and trimers in the derivations (section 4.2) are altered to S_1 and S_3.

4.1 Introduction

Contact between mammalian epithelia and the outer membrane (OM) of Gram-negative bacteria has an appreciable impact on symbiosis with these organisms (Koli, Sudan et al., 2011). The OM is 67% (w/w) outer membrane proteins (OMPs) (Jaroslawski, Duquesne et al., 2009). OMPs are insoluble, yet they must cross an aqueous

periplasm to reach the OM where they fold. Transport of unfolded OMPs (uOMPs) requires protection from self-aggregation by holding chaperones like the seventeen kilodalton protein (Skp) (Schäfer, Beck et al., 1999)

Skp possesses a fascinating propensity to bind unfolded proteins of diverse sequence (Burmann, Wang et al., 2013, De, Jeong et al., 2014). This binding underlies the Skp function as a holding chaperone (Walton, Sandoval et al., 2009) an activity accomplished by the enclosure of uOMPs within a cavity formed by a trimer of Skp monomers (Callon, Burmann et al., 2014). The trimer is thus accepted as the relevant species *in vivo* (Korndörfer, Dommel et al., 2004, Schlapschy, Dommel et al., 2004, Walton and Sousa, 2004, Stirling, Bakhoum et al., 2006, Qu, Mayer et al., 2007, Qu, Behrens-Kneip et al., 2009, Wu, Ge et al., 2011, Burmann and Hiller, 2012, Lyu, Shao et al., 2012, Burmann, Wang et al., 2013, McMorran, Bartlett et al., 2013, Patel and Kleinschmidt, 2013, Callon, Burmann et al., 2014). Evidence of this model includes binding experiments showing Skp saturates at a 3:1 Skp:uOMP ratio (Qu, Mayer et al., 2007), SANS analysis showing Skp−uOMP complexes form a 3:1 complex (Zaccai, Sandlin et al., 2016), and crystallographic analysis demonstrating that apo-Skp is trimeric (Walton and Sousa, 2004). Moreover, monomeric Skp has been observed only in a denatured form both on sodium dodecyl sulfate−polyacrylamide gel electrophoresis gels and by nuclear magnetic resonance (Schlapschy, Dommel et al., 2004). As a result, consideration of a role for monomeric Skp *in vivo* or *in vitro* has been minimal. However, concentrations of Skp used in structural studies are 10−100 times higher than the *in vivo* concentrations as estimated by quantitative LC/MS−MS (Korndörfer, Dommel et al., 2004, Schlapschy, Dommel et al., 2004, Masuda, Saito et al., 2009, Walton, Sandoval et

al., 2009, Arike, Valgepea et al., 2012, Burmann and Hiller, 2012, Burmann, Wang et al., 2013). This raises the question of whether Skp trimers exist at the lower concentrations more relevant to the cellular condition. To address this question, we used sedimentation equilibrium analytical ultracentrifugation (SE) to measure the mean trimerization constant for Skp, which we will term L_{13} (section 4.2, equations 4.10 and 4.11). Because ionic phospholipid headgroups are thought to interact with Skp (De Cock, Schäfer et al., 1999), electrostatically induced dissociation of Skp trimers could potentially be a mechanism by which Skp releases its uOMP cargo at the OM. Therefore, we conducted our experiments considering previous work suggesting that electrostatics affects Skp function. Qu et al. observed a weakening of the Skp:uOMP K_d in 1 M NaCl relying upon an analysis that assumed obligate Skp trimers (Qu, Mayer et al., 2007). Alternatively, a weakened L_{13} could explain this result. More recently, it was proposed that Skp monomers are unfolded yet fold and are stabilized by salt bridges upon trimer formation (Burmann, Holdbrook et al., 2015). On the other hand, the Skp structure (Korndörfer, Dommel et al., 2004) possesses five Asp pairs <7 Å apart (figure 4.3 and table 4.1), implying L_{13} could increase if ions screen this repulsion. Consistent with this view, De Cock et al. showed that $MgCl_2$ protects Skp from proteolysis, a possible consequence of an increased L_{13} (De Cock, Schäfer et al., 1999). Thus, the effect of salts on L_{13} is unclear. Using SE, we show here that Skp is not an obligate trimer at physiological concentrations, and we demonstrate that trimerization is independent of concentrations of several salts as well as large heat capacity changes (ΔC_p) over the ranges tested.

4.2 Materials and Methods

Protein Expression and Purification

We introduce here the nomenclature NZ100 to define the pOPINe construct containing Skp with a 6-histidine C-terminal tag (Skp-His) that was cloned as described in (Moon, Zaccai et al., 2013) and is used for all experiments. BL21 (DE3) pLysS cells in glycerol stock containing NZ100 were used to inoculate sterile LB with 60 μg/mL ampicillin, 34 μg/mL chloramphenicol, and 1% glucose. Overnight cultures grown at 37 °C were diluted 1:50 in 2XYT in baffled flasks and grown in the presence of antibiotic and glucose for 1.5 hours before changing to ambient temperature (RT). At 2 hours post inoculation bacteria were induced with 1 mM IPTG and grown overnight at RT. Cells were pelleted and frozen at -20 °C until use.

To purify Skp-His, cell pellets were placed on ice and resuspended in cold nickel buffer A (50 mM Tris-HCl (Sigma), 25 mM imidazole (Sigma), 0.5 M NaCl (Sigma), pH 8.0) containing EDTA-free Complete Protease Inhibitor (Roche) according to the manufacturer's instructions. Cell lysate was homogenized 2-3 times in an Emulsiflex homogenizer (Avestin) while kept on ice, and centrifuged at 19,000 rpm (41,241 g) in a Beckman J2-MI centrifuge in a JA-21 rotor. The resulting supernatant was filtered with a Millex GV 0.22 μm filter before binding to nickel-sepharose beads (GE Healthcare). The bench top column was washed with 10 column volumes of nickel buffer A and eluted with a 1:1 mixture of nickel buffer A and nickel buffer B (50 mM Tris, 500 mM imidazole, 0.5 M NaCl, pH 8.0). This eluent was further diluted with gel filtration buffer (GF, 20 mM Tris, 200 mM NaCl, pH 8.0) to an OD280 of 0.5 and then filtered. This Skp-His sample was transferred to GF buffer using a Bio-logic Duoflow FPLC (Bio-Rad)

utilizing a Superdex 200 10/300 GL column (GE Healthcare). The resultant fractions in GF at OD280 0.1 (67 μM) were concentrated to OD 0.5-0.65 with an Amicon filter (Millipore) with a 30 kD cutoff and stored at 4 °C at a concentration of approximately 250 μM until measurement. Stock concentrations were determined using the theoretical extinction coefficient ε_{280} =1490 M^{-1} cm^{-1} (Gill and von Hippel).

As a control for the effect of the His tag, Wild-type Skp was produced from a Skp construct with a 6-Histidine N-terminal tag (His-Skp) followed by a Tobacco Etch Virus cleavage site (Liu and Sturtevant). This construct is contained in a pET28b vector with kanamycin resistance in HMS (DE3) cells, and was expressed and purified in a similar fashion as Skp-His from NZ100, except for the following changes. Growths contained kanamycin (50 μg/mL) and no glucose, and cells were grown at 37 °C for 5 hours after induction with IPTG before pelleting cells. Elution of His-Skp from nickel-sepharose beads was accomplished with a TEV cleavage buffer (50 mM Tris, 0.5 M NaCl, 300 mM imidazole, 10 mM DTT, 2 mM EDTA, pH 8.0). To remove the N-terminal His tag, His-Skp in cleavage buffer was dialyzed overnight into nickel buffer A plus 10 mM DTT and 2 mM EDTA (pH 8.0) along with TEV at a 1:10 w:w ratio of TEV:His-Skp. The dialyzed cleavage reaction was added to nickel-sepharose beads as described above for Skp-His, and cleaved Skp (WT Skp) was recovered in the flow-through. WT Skp was then subjected to gel filtration to exchange the elution buffer for GF buffer in the same manner as Skp-His.

Purity of the samples, which we estimate as higher than 95%, was assessed by 5-20% gradient SDS-PAGE using a Bio-Rad Mini-Protean system at ambient temperature with settings of 200 volts for 28 minutes. Protein loads were 1 μg per lane (see figure 4.8).

Sedimentation Equilibrium

Skp-His samples in GF were diluted into 20 mM Tris buffer containing varying concentrations of the salts described in order to reach the final concentrations employed in the experiments (see figure captions). Samples with an OD230 = 0.90, 0.60, and 0.30 at a path length of 1.2 cm (corresponding to concentrations of 20, 13, and 7 μM) were loaded into six-sector centerpieces and allowed to reach equilibrium at speeds of 20,000, 24,500, and 30,000 rpm in a Beckman Optima XL-A analytical ultracentrifuge at the indicated temperatures. Data were recorded at 230 nm. The progress towards equilibrium was monitored with WINMATCH v0.99, and data editing and global fitting were performed, respectively, with WINREEDIT v0.999.0028 and WINNONLIN v1.06 (Johnson, Correia et al., 1981). The equilibrium data were edited resulting in greater than 800 degrees of freedom in most cases (see table 4.3, last column). The reduced molecular weight, sigma (s) as defined by Yphantis (Yphantis, 1964), partial specific volumes and buffer densities were calculated from the predicted Skp-His amino acid composition using SEDNTERP for each experimental condition at 20,000 rpm (Hayes, Laue et al., 1995). These are provided in table 4.5. Monomer molecular weights of 16645.9 Da and 16017.2 for Skp-His and WT Skp, respectively, were also calculated using SEDNTERP.

WINNONLIN reports the fitted parameter $\ln(K_{Abs})$, where K_{Abs} is the equilibrium constant in units of Absorbance (Gill and von Hippel, 1989). The following equation was used to convert $\ln(K_{Abs})$ into L_{13}, which has units of M^{-2}:

$$L_{13} = \frac{e^{\ln(K_{Abs})}\varepsilon^2 l^2}{3} \qquad \text{Equation 4.1}$$

where $l = 1.2$ cm and equals the path length, and ε is the molar extinction coefficient, which equals 37700 M^{-1} cm^{-1} and 35900 M^{-1} cm^{-1} for Skp-His and WT Skp,

respectively, at 230 nm. These values were determined using a standard curve on a Nanodrop spectrophotometer (Thermo Scientific). The reported mean and standard error of the mean (Stock, Rauch et al.) values for L_{13} are calculated from a minimum of three independent experiments. Data plotting was performed with Igor Pro 6.35A (Wavemetrics, Inc.).

To control for any possible effects due to the presence of the C-terminal six-His tag in Skp-His, we performed a set of experiments at 26 °C and all seven NaCl concentrations (see main text) using WT Skp. Figure 4.9 shows that the data for WT Skp are within the reproducibility of energies recorded for three replicates for Skp-His across all seven NaCl concentrations. Thus, we conclude that the C-terminal 6-His tag in Skp-His has minimal if any effect on trimerization or its sensitivity to salts.

Calculation of Ionic Activities

Ionic activities ($\alpha_{\pm}$) as reported in the x-axes of our figures and in table 4.7 were calculated using the basic Debye-Hückel approximation after Hamer (Hamer and Wu, 1972) for ions with formula X_aY_b:

$$\alpha_{\pm} = a^a b^b \gamma_{\pm}^{(a+b)} m^{(a+b)} \qquad \text{Equation 4.2}$$

where *m* is the molality and $\gamma_{\pm}$ is the mean ionic activity coefficient, defined as

$$\gamma_{\pm} \equiv \left(\gamma_{+}^{a}\gamma_{-}^{b}\right)^{\frac{1}{a+b}} \qquad \text{Equation 4.3}$$

and by the primitive Debye-Hückel approximation

$$\log(\gamma_{\pm}) = -|z_{+}||z_{-}|\frac{A\sqrt{I}}{1+Ba_0\sqrt{I}} \qquad \text{Equation 4.4}$$

A and B are functions of temperature and the dielectric constant of water, with simplified values

$$A = \frac{1.824829238 \times 10^6}{(T\epsilon)^{\frac{3}{2}}} \quad \text{Equation 4.5}$$

$$B = \frac{50.29158649 \times 10^8}{(T\epsilon)^{\frac{1}{2}}} \quad \text{Equation 4.6}$$

The parameter a_0 is a size parameter arbitrarily defined as

$$a_0 = 1 \times 10^{-11} \quad \text{Equation 4.7}$$

I is the ionic strength, defined as half the sum of the squares of the charges of each ion in the formula, weighted by their molalities

$$\mathrm{I} = \frac{1}{2}\sum_i \mathrm{z_i^2 m_i} \quad \text{Equation 4.8}$$

z_i is the charge on ion i, m_i is the molality of ion i, e is the dielectric constant for water, and T is the temperature in degrees Kelvin. Malberg and Maryott (Malmberg and Maryott, 1956) determined the empirical dependence of ϵ as

$$\epsilon = 87.740 - 0.40008\mathrm{t} + 9.398{\times}10^{-4}\mathrm{t}^2 - 1.410{\times}10^{-6}\mathrm{t}^3 \quad \text{Equation 4.9}$$

Where t is the temperature in degrees Celsius. This temperature dependence is accounted for in our figures.

Calculations for Figure 4.1A: Derivation of an Isotherm for Species Populations involved in Monomer-Trimer Association

The fraction of trimeric protein as a function of the total protein concentration can be expressed using the known monomer ↔ trimer equilibrium constant as follows. The reaction

$$3\mathrm{S}_1 \overset{\mathrm{L}_{13}}{\leftrightarrow} \mathrm{S}_3 \quad \text{Equation 4.10}$$

where the association equilibrium constant for Skp monomers to form trimers L_{13} can be written as:

$$L_{13} = \frac{S_3}{S_1^3}$$ Equation 4.11

where S_3 and S_1 are the molar concentrations of Skp trimers and free Skp monomers, respectively, and L_{13} has units of M^{-2}. Using the equation for conservation of mass,

$$S_0 = 3S_3 + S_1$$ Equation 4.12

an expression for the concentration of Skp trimer can be written as a function of total Skp (S_0), where S_0 is in monomeric units, by substituting (4.12) into (4.11) as follows:

$$S_3 = L_{13}(S_0 - 3S_3)^3$$ Equation 4.13

Expansion and collection of terms leads to the following third order polynomial equation in S_3:

$$S_3^3 - S_0 S_3^2 + \left(\frac{9S_0^2 + L_{13}^{-1}}{27}\right) S_3 - \frac{S_0^3}{27} = 0$$ Equation 4.14

The solution to this polynomial has a single real root that gives the trimer concentration as a function of total Skp, which equals:

$$S_3(S_0) = \frac{S_0}{3} + \frac{\xi}{\left(\sqrt{\beta^2 - \xi^3} + \beta\right)^{\frac{1}{3}}} + \left(\sqrt{\beta^2 - \xi^3} + \beta\right)^{\frac{1}{3}}$$ Equation 4.15

where δ, ξ, and β are defined as follows:

$$\delta = \frac{9L_{13}S_0^2 + 1}{L_{13}}$$ Equation 4.16

$$\xi = \frac{S_0^2}{9} - \frac{\delta}{81}$$ Equation 4.17

$$\beta = \frac{S_0^3}{18} - \frac{S_0 \delta}{162}$$ Equation 4.18

The fraction of total Skp protein that is trimeric at any total Skp concentration equals:

$$f_{S_3} = \frac{3S_3}{S_0}$$ Equation 4.19

and the fraction of total Skp protein that is monomeric equals:

$$f_{S_1} = 1 - f_{S_3}$$ Equation 4.20

These fractions of Skp are plotted in figure 1A, using the best-fit value of $L_{13} = 6.18 \times 10^{11}$ M^{-2} for the 37 °C, 150 mM NaCl condition (table 4.5).

Calculations of ΔC_p Due to Hydration Effects for Folding (ΔC_{fold}) and Binding (ΔC_{13})

The equation

$$\Delta C_p^{hydration} = C_n \Delta A_n + C_p \Delta A_p$$ Equation 4.21

represents an empirical estimate for ΔC_p due to hydration effects. The nomenclature in equation 21 follows Prabhu and Sharp, where C_n and C_p to the right of the equals sign are empirical parameters and ΔA_n and ΔA_p are changes in solvent-accessible surface areas between two states of interest calculated from structures (Prabhu and Sharp, 2005). Note that the subscripts n and p in this context only stand for "nonpolar" and "polar". (The subscript p in other contexts refers to constant pressure (e.g. ΔC_p, stands for "heat capacity change at constant pressure".) We use the values of $C_n = 0.66 \pm 0.21$ J mol^{-1} $^{\circ}K^{-1}$ $Å^{-2}$and $C_p = 0.52 \pm 0.32$ J mol^{-1} K^{-1} $Å^{-2}$ provided by Robertson and Murphy from multiple linear regression of their dataset of 49 proteins (Robertson and Murphy, 1997).

We used the POPS algorithm (Cavallo, Kleinjung et al., 2003) to calculate the nonpolar and polar solvent-accessible-surface-areas (SASAs) from PDB files of each of the Skp conformations. The appropriate subtractions lead to the values of ΔA_n and ΔA_p for a reaction of interest. The unfolded Skp-His model was calculated following the protocol of Robertson and Murphy where the SASA values of the "X" residue in an Ala-

X-Ala tripeptide were summed weighted according to the amino acid composition. For folded Skp-His monomers and trimers, we started with a model of Skp-His relaxed using molecular dynamics as described previously (Zaccai, Sandlin et al., 2016).

The $\Delta C_p^{hydration}$ for the unfolding to folding structural change of each monomer (ΔC_{fold}) equals:

$$\Delta C_{Fold} = C_n \Delta A_{n,Folded-Unfolded} + C_p \Delta A_{p,Folded-Unfolded} \quad \text{Equation 4.22}$$

These values are calculated for each of the three monomers and summed to calculate the value for the trimer.

The $\Delta C_p^{hydration}$ for the trimerization of three folded monomers (ΔC_{13}) equals:

$$\Delta C_{13} = C_n \Delta A_{n,Trimer-Sum3FoldedMonomers} + C_p \Delta A_{p,Trimer-Sum3FoldedMonomers} \quad \text{Equation 4.23}$$

The total ΔC_p for folding of three monomers and trimerization equals the sum of ΔC_{Fold} for each monomer and ΔC_{13}. Table 4.6 shows a summary of these calculations.

Fitting of Temperature-dependence SE Data to the Integrated van't Hoff Equation, Associated Jackknife ("Bootstrap") Statistics, and Comparison With Fitting to the Linearized Form of the Equation

Fit to Integrated van't Hoff Equation

In response to a reviewer's suggestion to provide quantitative estimates of small heat capacity changes that occur upon trimerization of Skp, we fit the temperature-dependence data at seven NaCl concentrations to a form of the integrated van't Hoff equation (Liu and Sturtevant, 1995) using Igor Pro 6.35A:

$$\ln(L_{13}) = \frac{1}{R}\left[\left(\Delta H^{\circ}_{ref} - T_{ref}\Delta C_{p,T_{ref}}\right)\left(\frac{1}{T_{ref}} - \frac{1}{T}\right) - \frac{1}{R}\Delta C_{p,T_{ref}} \ln\left(\frac{T_{ref}}{T}\right)\right] + \ln L_{13,T_{ref}}$$

Equation 4.24

where T_{Ref} is assumed to be 37 ˚C (310.15 ˚K) for this fitting, $\Delta H^{\circ}_{37\,^{\circ}C}$ and $\Delta C_{p,37\,^{\circ}C}$ are the floating parameters whose values are estimated from the fits, and L_{13} at 37 ˚C is fixed to the experimentally observed value.

Bootstrap Procedure

To obtain error estimates of the thermodynamic parameters, we used a bootstrapping approach (commonly referred to as a jackknife) in which each NaCl dataset (50 mM, 150 mM, 200 mM, 250 mM, 350 mM, 600 mM, and 1000 mM) was fit six times. In each fit, one of the six data points was removed and the best-fit parameters from the resultant regression were recorded. The mean and standard error of the resulting six best-fit parameters are the reported statistics. These values are shown in figure 4.5A. Confidence lines for fits to the integrated van't Hoff equation shown in figure 4.5A were computed by first fixing the enthalpy change to the bootstrapped value and then simulating the curves deriving from the bootstrapped ΔC_p ± bootstrapped error.

Fit to Linearized van't Hoff Equation

Initially, we analyzed our data using the simpler linear form of the van't Hoff equation:

$$\ln(L_{13}) = -\frac{\Delta H^{\circ}}{R}\frac{1}{T} + \frac{\Delta S^{\circ}}{R} \qquad \text{Equation 4.25}$$

where $-\frac{\Delta H^{\circ}}{R}$ and $\frac{\Delta S^{\circ}}{R}$ are the slope and intercept, respectively, estimated in the linear regression (see figure 4.5C). After comparing this model to the integrated van't Hoff equation with fixed $\ln(L_{13,37\,^{\circ}C})$, which possesses the same number of floating parameters,

we chose the integrated equation for three reasons. First, the integrated equation is a more physically correct model in the sense that it explicitly accounts for a heat capacity change and temperature dependence of the enthalpy change, however small their magnitudes. Second, the residuals across all seven NaCl concentrations appear systematic in the linear fit in a way that is consistent with limited curvature. Third, the sums of squared residuals (SSR) for each fit are statistically better for the integrated equation.

4.3 Results

All self-association experiments were conducted using SE. Figure 1B and figure 4.4A−J and table 4.4 show that a monomer−trimer association model best described the data. In contrast, models including single species or other self-association schemes showed larger normalized error and systematic residuals. We therefore analyzed our data using the simpler monomer−trimer model for self-association. The mean trimerization constant (L_{13}) from three replicates for the 37 °C, 150 mM NaCl, pH 8.0 condition was 6.18 x 10^{11} M^{-2}, which corresponds to a $\Delta G°$ of −16.73 ± 0.29 kcal mol^{-1} . Figure 4.1A shows this predicts a concentration ($C_{1/2}$) of 1.47 μM at which half of Skp is trimeric. Using values from the literature for Skp copy numbers (Masuda, Saito et al., 2009), and an envelope volume of 0.14 μm^3 (Goodsell, 1991), the concentration of Skp in the periplasm is approximately 3.9 μM under stationary phase growth in LB and 31.0 μM under steady-state growth in M9/glucose medium (figure 4.1A). These considerations result in monomeric fractions of 29 and 8%, respectively, for the two growth conditions (figure 4.1A). Thus, trimers represent only part of the Skp population in *E. coli* when cells are not nutritionally stressed. Under stress conditions, the fraction of trimeric Skp is

predicted to increase. Ions in either the membrane interface or periplasm could affect this fraction by serving as local triggers for complex dissociation or formation. To test this idea, we measured L_{13} at seven NaCl concentrations (50, 150, 200, 250, 350, 600, and 1000 mM) (figure 4.2A and table 4.5). We also held NaCl constant (50 mM) and varied $MgCl_2$ (1, 2, and 10 mM) and Na_3PO_4 (1, 10, and 100 mM) (figure 4.2B and table 4.5). figure 4.2 shows that L_{13} does not depend on salt concentration; this indicates a negligible role for long-range electrostatics in Skp trimerization. Burmann et al. hypothesized that salt bridges stabilize the fold of Skp subunits (Burmann, Holdbrook et al., 2015). This model invokes a linkage among salt concentration, folding, and trimerization. To address whether folding occurs upon trimerization, and because protein folding results in a ΔC_p that can be observed as curvature in a van't Hoff plot, we analyzed temperature-dependent SE data. Figure 4.2C and figure 4.5A show that a fit of $\ln L_{13}$ versus $1/T$ to the integrated van't Hoff equation returns a ΔC_p of -0.62 ± 0.11 kcal mol^{-1} K^{-1} for the 150 mM NaCl condition. To compare our experimental ΔC_p to that expected, we calculated the ΔC_p for folding (ΔC_{fold}) and trimerization (ΔC_{13}) (see section 4.2 and table 4.6) and obtained estimates for ΔC_{fold} of -7.7 ± 3.3 kcal mol^{-1} K^{-1} for the folding of three monomers and for ΔC_{13} of -0.31 ± 0.12 kcal mol^{-1} K^{-1} for trimerization of three folded monomers, which sum to a $\Delta C_{p,\,total}$ of -8.01 ± 3.3 kcal mol^{-1} K^{-1} for both reactions. Our analysis thus indicates that trimerization occurs between folded monomers. We also found that ΔC_p is largely insensitive to NaCl concentration (figure 4.7). We cannot rule out a very small amount of partial unfolding, but overall, we find no evidence of folding linked to trimerization that is stabilized by salts. Instead, we observe enthalpy−entropy compensation (figure 4.6).

4.4 Discussion

Electrostatics has been connected with Skp since it was first copurified with LPS (Geyer, Galanos et al., 1979), inspiring many studies (Qu, 2007), and with DNA, which produced the obsolete alias HLP-I (histone-like protein I) (Holck and Kleppe, 1988). Even so, our results show no linkage between L_{13} and physiological ion activities.

Because Qu et al. observed an increased K_d for uOmp binding at 1 M NaCl, they suggested Coulombic forces dominate uOMP binding. We ruled out an alternative hypothesis, that L_{13} weakens in 1 M NaCl. However, the insensitivity of Skp trimerization to NaCl suggests a third possibility: that uOMP aggregation is accelerated in high salt (Ebie Tan, Burgess et al., 2010). Such a situation would lead to a depletion of monomeric uOMP and an increased apparent K_d. Moreover, while $MgCl_2$ protects Skp from proteolysis (De Cock, Schäfer et al., 1999), our results suggest this was likely not due to an increased L_{13}.

The existence of a significant population of monomeric Skp requires a reexamination of results from both in vitro and in vivo experiments. At the highest concentrations of Skp used for in vitro fluorescence binding assays (Qu, Mayer et al., 2007), more than 50% of Skp is monomeric, yet after equilibrium is reached, a 3:1 Skp:uOMP stoichiometry is observed. This is a consequence of thermodynamic linkage: binding of uOMP substrates stabilizes the quaternary Skp:uOMP complex. Importantly, descriptions of existing (Wu, Ge et al., 2011) and future kinetics data for complex assembly require models that account for the monomeric state. We suggest that this

significant population of monomers may increase the diversity of species that could assemble around a client.

Figure 4.1 shows that the fraction of Skp trimers may be regulated by cellular stress (Tao, Bausch et al., 1999). If trimers possess the bulk of holding activity (Entzminger, Chang et al., 2012), then regulation of Skp concentration regulates this activity. Such regulation of functional multimers by gene expression has been observed before (Cai, Lee et al., 1990). Activation of Skp under stress conditions is also consistent with a model in which Skp delivers uOMPs to DegP for degradation (Wu, Ge et al., 2011). A similar model was proposed for the structurally and functionally analogous cochaperones Prefoldin and CCT in eukaryotes (Lundin, Leroux et al., 2010).

In summary, our results support a view that Skp is more than a static cage. Instead, the protein experiences a dynamic monomer−trimer equilibrium and has evolved to respond to changing growth environments while remaining robust to commensurate changes in salt concentration.

4.5 Tables and Figures

Table 4.1. Aspartate residues in the Skp trimerization interface within distances where Coulombic forces may be significant.

Involved Residues	Involved Chains	Range (Å)
Asp14 - Asp140	A-B, C-A	5.8, 4.2
Asp126 - Asp140	A-B, B-C, A-C	5.1, 6.5, 5.4

Table 4.2. A comparison of normalized least-squares square root of variance (SRV) for five tested models for the 37 °C, 150 mM NaCl, pH 8.0 condition shows that the Monomer-Trimer association scheme is the best fit to the data.

Replicate Number[a]	Obligate Monomer SRV/ DoF[b]	Obligate Trimer SRV/ DoF [b]	Monomer-Trimer SRV/ DoF [b]	Monomer-Dimer-Trimer SRV/ DoF [b]	Dimer-Tetramer SRV/ DoF [b]	DoF
1	3.49E-5	1.53E-5	5.87E-6	3.33E-5	6.87E-6	867
2	5.01E-5	1.09E-5	5.97E-6	5.90E-6	6.18E-6	803
3	4.42E-5	7.11E-6	5.15E-6	7.15E-5	5.08E-6	910
Mean (SRV/DoF)	4.31E-5	1.11E-5	5.66E-6	3.69E-5	6.04E-6	

[a.] The data for the first replicate are shown in figures 4.4A-J.
[b.] The square root of the variance (SRV) reported by WINNONLIN divided by the number of degrees of freedom for the global fit (DoF).

Table 4.3. Thermodynamic Parameters for Skp Trimerization.

T (°C)	NaCl (M)	Na_3PO_4 (M)	MgCl (M)	Mean L_{13} (M^{-2})	SEM L_{13} (M^{-2})	ΔG° ± SD (kcal mol^{-1})	Mean SRV	Mean DoF
4	0.050	0.000	0.000	1.03E13	6.76E12	-16.50 ± 0.28	4.71E-03	891
4	0.150	0.000	0.000	2.63E12	1.07E12	-15.75 ± 0.19	4.40E-03	861
4	0.200	0.000	0.000	4.09E12	1.46E12	-15.99 ± 0.17	4.33E-03	849
4	0.250	0.000	0.000	1.43E13	1.11E13	-16.68 ±0.32	4.15E-03	826
4	0.350	0.000	0.000	4.87E12	3.93E12	-16.09 ±0.33	4.64E-03	831
4	0.600	0.000	0.000	4.87E12	1.35E12	-16.09 ±0.13	4.50E-03	830
4	1.000	0.000	0.000	2.96E12	1.39E12	-15.81 ±0.21	4.37E-03	828
15	0.050	0.000	0.000	1.19E14	1.16E14	-18.56 ±0.37	4.46E-03	855
15	0.150	0.000	0.000	4.79E12	3.11E12	-16.72 ±0.28	4.37E-03	861
15	0.200	0.000	0.000	3.77E12	2.60E12	-16.58 ±0.29	4.41E-03	863
15	0.250	0.000	0.000	7.85E12	6.18E12	-17.00 ±0.32	4.36E-03	821
15	0.350	0.000	0.000	3.76E12	3.11E12	-16.58 ±0.33	4.56E-03	842
15	0.600	0.000	0.000	3.81E12	1.62E12	-16.59 ±0.19	4.64E-03	848
15	1.000	0.000	0.000	1.70E12	4.91E11	-16.12 ±0.14	4.27E-03	791
19	0.050	0.000	0.000	1.19E13	9.61E12	-17.48 ±0.33	4.92E-03	864
19	0.150	0.000	0.000	2.44E12	1.42E12	-16.56 ±0.25	4.44E-03	854
19	0.200	0.000	0.000	1.97E13	1.85E13	-17.77 ±0.36	4.60E-03	829
19	0.250	0.000	0.000	2.86E12	1.36E12	-16.65 ±0.21	4.61E-03	779
19	0.350	0.000	0.000	2.44E12	1.39E12	-16.56 ±0.25	4.61E-03	824
19	0.600	0.000	0.000	2.72E12	1.77E12	-16.62 ±0.28	4.64E-03	830
19	1.000	0.000	0.000	1.02E12	3.01E11	-16.05 ±0.14	4.53E-03	734
22	0.050	0.000	0.000	7.01E12	4.31E12	-17.35 ±0.26	4.44E-03	850
22	0.150	0.000	0.000	1.91E12	5.61E11	-16.59 ±0.14	4.34E-03	798
22	0.200	0.000	0.000	4.83E12	3.51E12	-17.13 ±0.30	4.60E-03	843
22	0.250	0.000	0.000	1.85E12	5.06E11	-16.57 ±0.13	4.57E-03	794
22	0.350	0.000	0.000	2.12E12	1.09E12	-16.65 ±0.23	4.67E-03	822
22	0.600	0.000	0.000	2.95E12	9.26E11	-16.84 ±0.15	4.48E-03	827
22	1.000	0.000	0.000	1.08E12	2.40E11	-16.25 ±0.11	4.53E-03	737
26	0.050	0.000	0.000	1.61E13	1.49E13	-18.07 ±0.36	4.52E-03	876
26	0.150	0.000	0.000	1.49E12	1.06E12	-16.66 ±0.30	4.57E-03	832
26	0.200	0.000	0.000	1.13E12	7.29E11	-16.49 ±0.27	4.15E-03	864
26	0.250	0.000	0.000	2.89E12	2.45E12	-17.05 ±0.34	4.11E-03	826
26	0.350	0.000	0.000	1.33E12	1.02E12	-16.59 ±0.31	4.35E-03	825
26	0.600	0.000	0.000	1.41E12	9.62E11	-16.63 ±0.29	4.53E-03	835
26	1.000	0.000	0.000	7.58E11	3.81E11	-16.26 ±0.22	4.22E-03	784

Table 4.3 (Continued). Thermodynamic Parameters for Skp Trimerization

T (°C)	NaCl (M)	Na_3PO_4 (M)	$MgCl_2$ (M)	L_{13} (M^{-2})	SEM L_{13}	DG° ± SEM (kcal mol^{-1})		Mean SRV	Mean DoF
37	0.050	0.000	0.000	1.60E12	1.11E12	-17.32	±0.29	4.40E-03	893
37	0.150	0.000	0.000	6.18E11	4.28E11	-16.73	±0.29	4.85E-03	859
37	0.200	0.000	0.000	5.47E11	3.04E11	-16.66	±0.24	4.47E-03	883
37	0.250	0.000	0.000	8.57E11	6.25E11	-16.93	±0.30	2.94E-03	839
37	0.350	0.000	0.000	8.38E11	6.10E11	-16.92	±0.30	4.61E-03	848
37	0.600	0.000	0.000	1.04E12	3.84E11	-17.05	±0.17	4.40E-03	866
37	1.000	0.000	0.000	5.91E11	3.47E11	-16.70	±0.25	4.35E-03	799
26	0.050	0.001	0.000	2.95E12	1.61E12	-16.62	±0.26	3.30E-03	685
26	0.050	0.010	0.000	3.47E12	8.23E11	-17.11	±0.13	2.91E-03	612
26	0.050	0.100	0.000	1.21E12	2.18E11	-16.51	±0.10	3.58E-03	592
26	0.050	0.000	0.001	5.03E12	4.58E12	-16.85	±0.38	2.94E-03	651
26	0.050	0.000	0.002	1.11E13	8.69E12	-17.17	±0.34	3.46E-03	691
26	0.050	0.000	0.010	5.62E12	2.93E12	-17.15	±0.25	3.02E-03	639

Table 4.4. POPS SASA calculations for Skp and Estimated ΔC_p

Chain	Conformation	Nonpolar ASA (Å^2)	Polar ASA (Å^2)	$DA_{n,F-U}$ (Å^2)[a]	$DA_{P,F-U}$ (Å^2)[b]	$DA_{n,T-3M}$ (Å^2)[d]	$DA_{p,T-3M}$ (Å^2)[e]	DC_{fold}[f]	DC_{13}[f]
uSkp-His	Extended	16730.9	12872.5	N/A	N/A	N/A	N/A	N/A	N/A
A	Folded Monomer	6520.3	5327.0	-8582.9	-3892.9	N/A	N/A	-2.5±1.1	N/A
B	Folded Monomer	6494.4	5335.7	-8608.8	-3884.1	N/A	N/A	-2.6±1.1	N/A
C	Folded Monomer	6483.0	5288.4	-8620.1	-3931.5	N/A	N/A	-2.6±1.1	N/A
A,B,C	Sum Three Folded Protomers	19497.8	15951.0	-25811.8	-11708.4	NA	NA	-7.7±3.3	NA
A:B:C	Trimer	18021.6	15314.9	N/A	N/A	-1476.2	-636.2	N/A	-0.31±0.12

[a] Change in nonpolar SASA for folding from extended conformational model.
[b] Change in polar SASA for folding from extended conformational model.
[d] Change in nonpolar SASA for binding of three folded monomers.
[e] Change in polar SASA for binding of three folded monomers.
[f] Units are kcal mol^{-1} K^{-1}

Table 4.5. Calculated Activities and Parameters Used for WINNONLIN Least-squares Regression

Construct	T (°C)	NaCl (M)	Na_3PO_4 (M)	$MgCl_2$ (M)	Activity (Unitless)[a]	$\bar{v}$ (mL g^{-1})	r (g mL^{-1})	Monomer s, 20k rpm (cm^{-2})
Skp-His	4	0.050	0.000	0.000	1.50E-03	0.72090	1.0027	0.87820
Skp-His	4	0.150	0.000	0.000	9.31E-03	0.72090	1.0068	0.86883
Skp-His	4	0.200	0.000	0.000	1.44E-02	0.72090	1.0088	0.86427
Skp-His	4	0.250	0.000	0.000	2.00E-02	0.72090	1.0109	0.85947
Skp-His	4	0.350	0.000	0.000	3.18E-02	0.72090	1.0150	0.85010
Skp-His	4	0.600	0.000	0.000	6.17E-02	0.72090	1.0251	0.82703
Skp-His	4	1.000	0.000	0.000	1.03E-01	0.72090	1.0409	0.79094
Skp-His	15	0.050	0.000	0.000	1.49E-03	0.72557	1.0018	0.83239
Skp-His	15	0.150	0.000	0.000	9.18E-03	0.72557	1.0059	0.82333
Skp-His	15	0.200	0.000	0.000	1.42E-02	0.72557	1.0079	0.81891
Skp-His	15	0.250	0.000	0.000	1.96E-02	0.72557	1.0100	0.81426
Skp-His	15	0.350	0.000	0.000	3.11E-02	0.72557	1.0141	0.80520
Skp-His	15	0.600	0.000	0.000	5.99E-02	0.72557	1.0242	0.78286
Skp-His	15	1.000	0.000	0.000	9.87E-02	0.72557	1.0400	0.74792
Skp-His	19	0.050	0.000	0.000	1.48E-03	0.72727	1.0011	0.81741
Skp-His	19	0.150	0.000	0.000	9.12E-03	0.72727	1.0052	0.80845
Skp-His	19	0.200	0.000	0.000	1.41E-02	0.72727	1.0072	0.80407
Skp-His	19	0.250	0.000	0.000	1.95E-02	0.72727	1.0093	0.79948
Skp-His	19	0.350	0.000	0.000	3.08E-02	0.72727	1.0134	0.79052
Skp-His	19	0.600	0.000	0.000	5.92E-02	0.72727	1.0235	0.76844
Skp-His	19	1.000	0.000	0.000	9.72E-02	0.72727	1.0393	0.73390
Skp-His	22	0.050	0.000	0.000	1.48E-03	0.72855	1.0005	0.80659
Skp-His	22	0.150	0.000	0.000	9.08E-03	0.72855	1.0046	0.79770
Skp-His	22	0.200	0.000	0.000	1.40E-02	0.72855	1.0066	0.79336
Skp-His	22	0.250	0.000	0.000	1.94E-02	0.72855	1.0087	0.78881
Skp-His	22	0.350	0.000	0.000	3.06E-02	0.72855	1.0128	0.77992
Skp-His	22	0.600	0.000	0.000	5.86E-02	0.72855	1.0229	0.75803
Skp-His	22	1.000	0.000	0.000	9.61E-02	0.72855	1.0386	0.72400
Skp-His	26	0.050	0.000	0.000	1.47E-03	0.73025	0.9948	0.79300
Skp-His	26	0.150	0.000	0.000	9.02E-03	0.73025	1.0036	0.78416
Skp-His	26	0.200	0.000	0.000	1.39E-02	0.73025	1.0056	0.77988
Skp-His	26	0.250	0.000	0.000	1.92E-02	0.73025	1.0077	0.77537
Skp-His	26	0.350	0.000	0.000	3.03E-02	0.73025	1.0118	0.76659
Skp-His	26	0.600	0.000	0.000	5.79E-02	0.73025	1.0218	0.74515
Skp-His	26	1.000	0.000	0.000	9.45E-02	0.73025	1.0376	0.71128
Skp-His	37	0.050	0.000	0.000	1.46E-03	0.73492	0.9960	0.75888

Table 4.5 (Continued). Calculated Activities and Parameters Used for WINNONLIN Least-squares Regression

Construct	T (°C)	NaCl (M)	Na_3PO_4 (M)	$MgCl_2$ (M)	Activity (Unitless)[a]	$\bar{v}$ (mL g^{-1})	r (g mL^{-1})	Monomer s 20k rpm (cm^{-2})
Skp-His	37	0.150	0.000	0.000	8.85E-03	0.73492	1.0001	0.75037
Skp-His	37	0.200	0.000	0.000	1.36E-02	0.73492	1.0021	0.74620
Skp-His	37	0.250	0.000	0.000	1.87E-02	0.73492	1.0042	0.74183
Skp-His	37	0.350	0.000	0.000	2.95E-02	0.73492	1.0083	0.73330
Skp-His	37	0.600	0.000	0.000	5.57E-02	0.73492	1.0183	0.71249
Skp-His	37	1.000	0.000	0.000	8.99E-02	0.73492	1.0340	0.67982
Skp-His	26	0.050	0.001	0.000	6.06E-07	0.73025	0.9997	0.79257
Skp-His	26	0.050	0.010	0.000	2.05E-05	0.73025	1.0015	0.78867
Skp-His	26	0.050	0.100	0.000	6.70E-05	0.73025	1.0182	0.75287
Skp-His	26	0.050	0.000	0.001	3.49E-09	0.73025	0.9996	0.79278
Skp-His	26	0.050	0.000	0.002	2.64E-08	0.73025	0.9996	0.79278
Skp-His	26	0.050	0.000	0.010	2.59E-06	0.73025	1.0001	0.79167
WT Skp	26	0.050	0.000	0.000	1.50E-03	0.73249	0.9948	0.75672
WT Skp	26	0.150	0.000	0.000	9.31E-03	0.73249	1.0036	0.74820
WT Skp	26	0.200	0.000	0.000	1.44E-02	0.73249	1.0056	0.74406
WT Skp	26	0.250	0.000	0.000	2.00E-02	0.73249	1.0077	0.73971
WT Skp	26	0.350	0.000	0.000	3.18E-02	0.73249	1.0118	0.73123
WT Skp	26	0.600	0.000	0.000	6.17E-02	0.73249	1.0218	0.71054
WT Skp	26	1.000	0.000	0.000	1.03E-01	0.73249	1.0376	0.67785

[a.] The activity plotted in logarithmic units in the x-axis of figures 2A and 2B for either NaCl (for conditions ranging from 50 mM to 1 M NaCl), or for $MgCl_2$ or Na_3PO_4 if these are present in the columns 4 or 5. Calculation of these activities is described in the SI methods section.

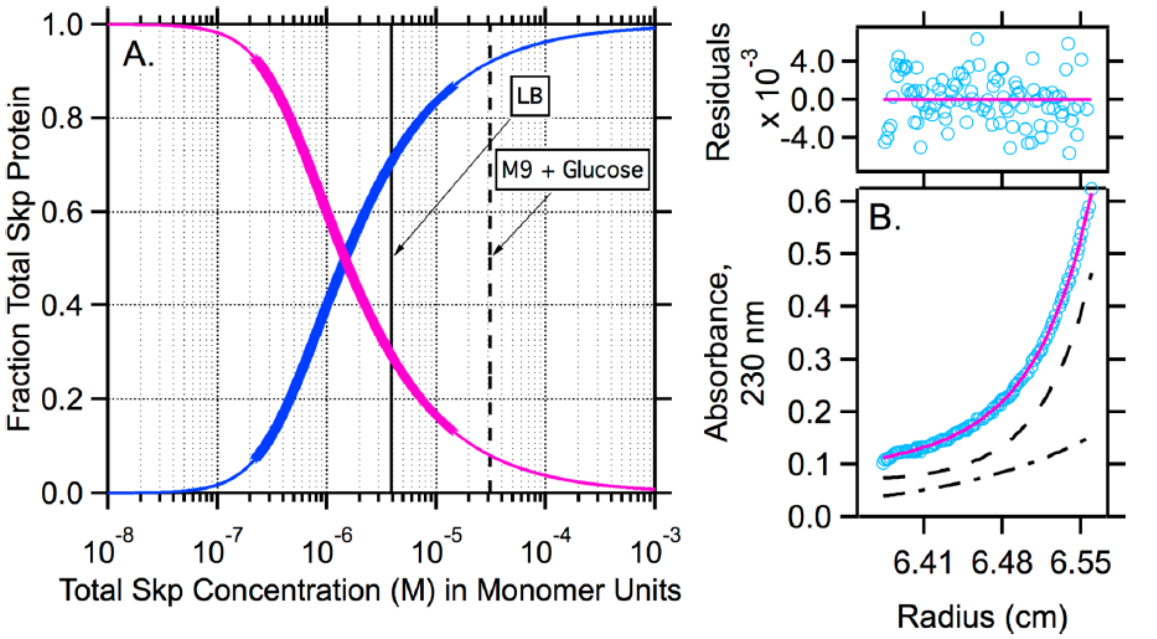


Figure 4.1. (A) Skp monomer populations are significant and depend on growth conditions. The fraction of total Skp in trimeric form is colored blue; the fraction of monomeric Skp is colored magenta. The thickened regions of the traces indicate the observable range in SE. This plot was calculated using the best fit of L_{13} = 6.18 Å x 10^{11} M^{-2} corresponding to 37 °C and 150 mM NaCl at pH 8.0, similar to the conditions in intestinal epithelia (Chatton and Spring, 1995). The solid black line denotes the estimated concentration of Skp in an Escherichia coli envelope during stationary phase growth in LB medium (Masuda, Saito et al.). The dashed black line denotes the Skp concentration as upregulated by sigmaE during steady-state growth in M9 glucose medium (Arike, Valgepea et al., 2012). (B) Species plot for SE data collected at 37 °C in 150 mM NaCl and 13 μM Skp at pH 8.0 with a rotor speed of 24500 rpm. The light blue circles denote A230. The magenta line is a fit to a monomer−trimer association model. Also shown are the contributing distributions of the trimeric (−−−) and monomeric species (·−·).

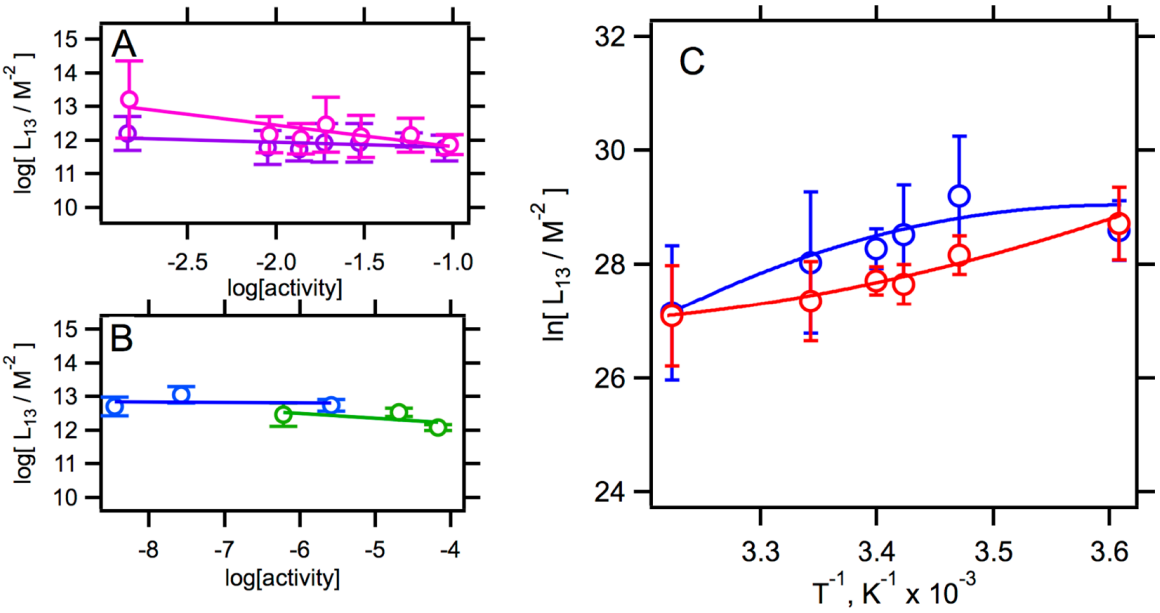


Figure 4.2. **Skp trimerization is not strongly affected by salt activity and displays small ΔC_p values.** (A) Linkage data for NaCl at 26 °C (magenta) and 37 °C (purple) are essentially independent of NaCl activity. (B) $MgCl_2$ activity (blue) and Na_3PO_4 activity (green) at 26 °C are not linked to Skp trimerization. The relationship between activity and concentration is described in section 5.2. (C) van't Hoff plot for Skp trimerization. Using fits to the integrated van't Hoff equation (see also section 5.2 and figure 4.5), Skp trimerization displays a small ΔC_p at both 150 mM (blue) and 1 M NaCl (red).

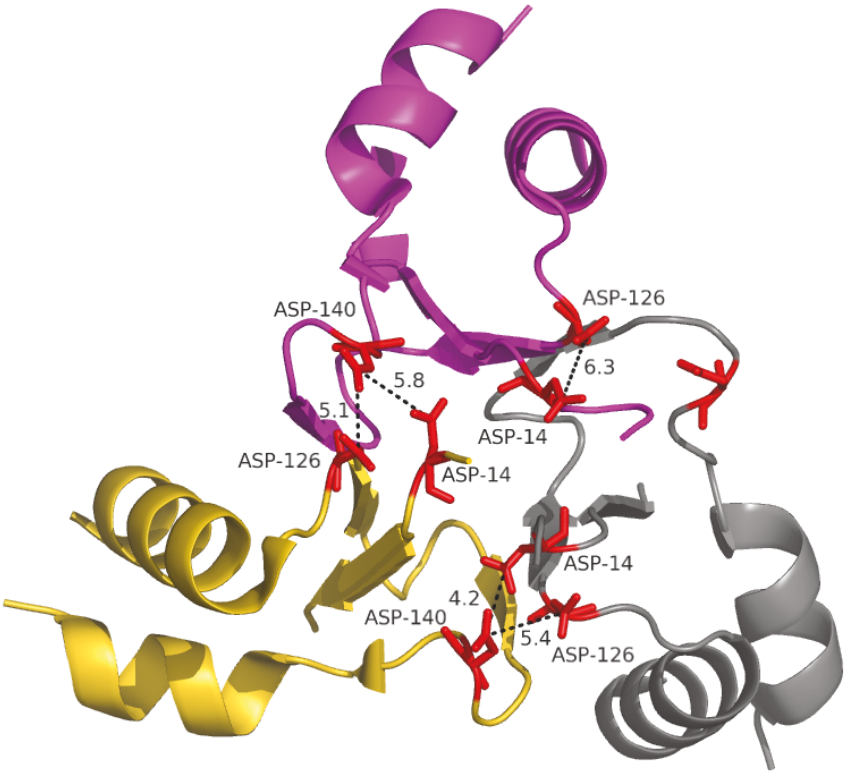


Figure 4.3. Potential Coulombic interactions across the Skp-Skp interface. Five aspartate residues identified in structure 1SG2 were found to be within inter-subunit distances (see table 4.3) at which Coulombic forces may be both significant and modulated by salts (Lee, Fitch et al., 2002).

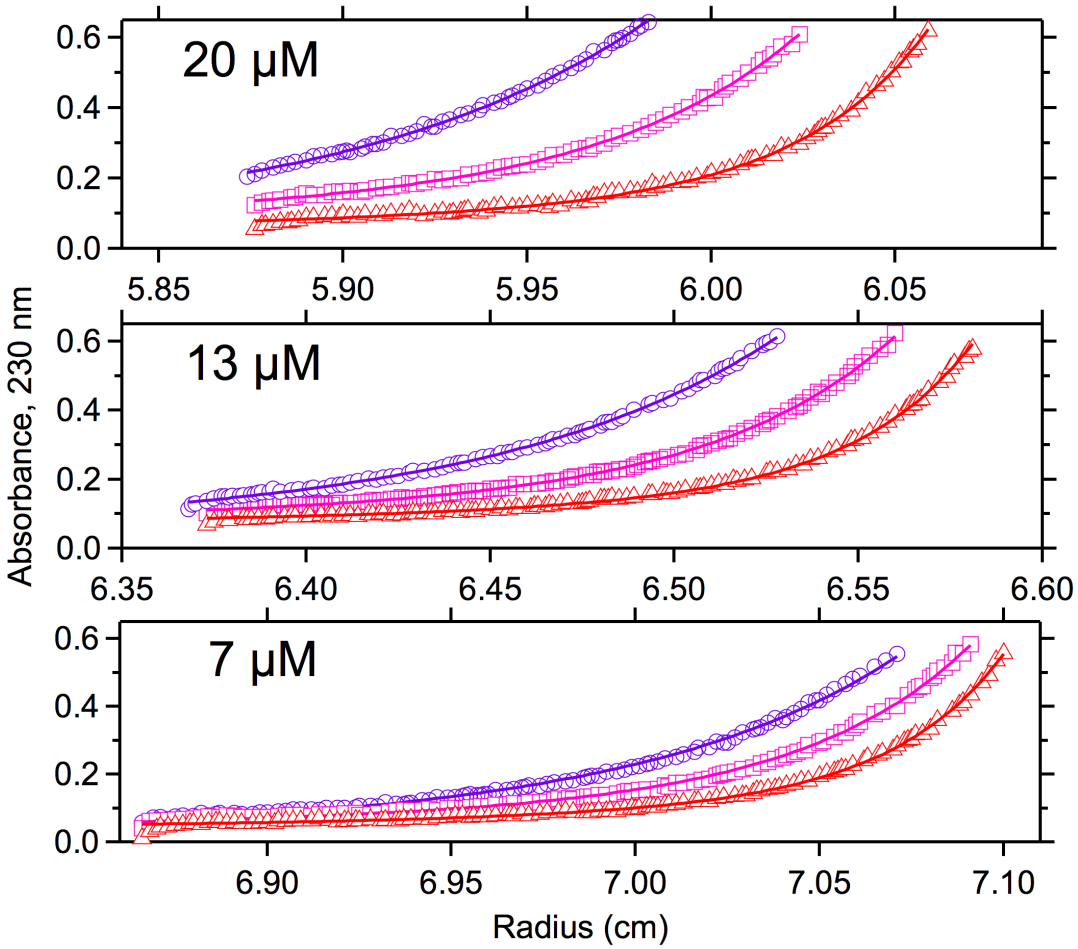


Figure 4.4A. A monomer-trimer association model describes the data. Representative processed Skp AUC SE data collected at 37 °C and 150 mM NaCl, pH 8.0 fit to a monomer-trimer association scheme. Symbols are as follows: purple circles, 20,000 rpm; magenta squares, 24,500 rpm; red triangles, 30,000 rpm. The residuals for this fit are shown in figure 4.4B.

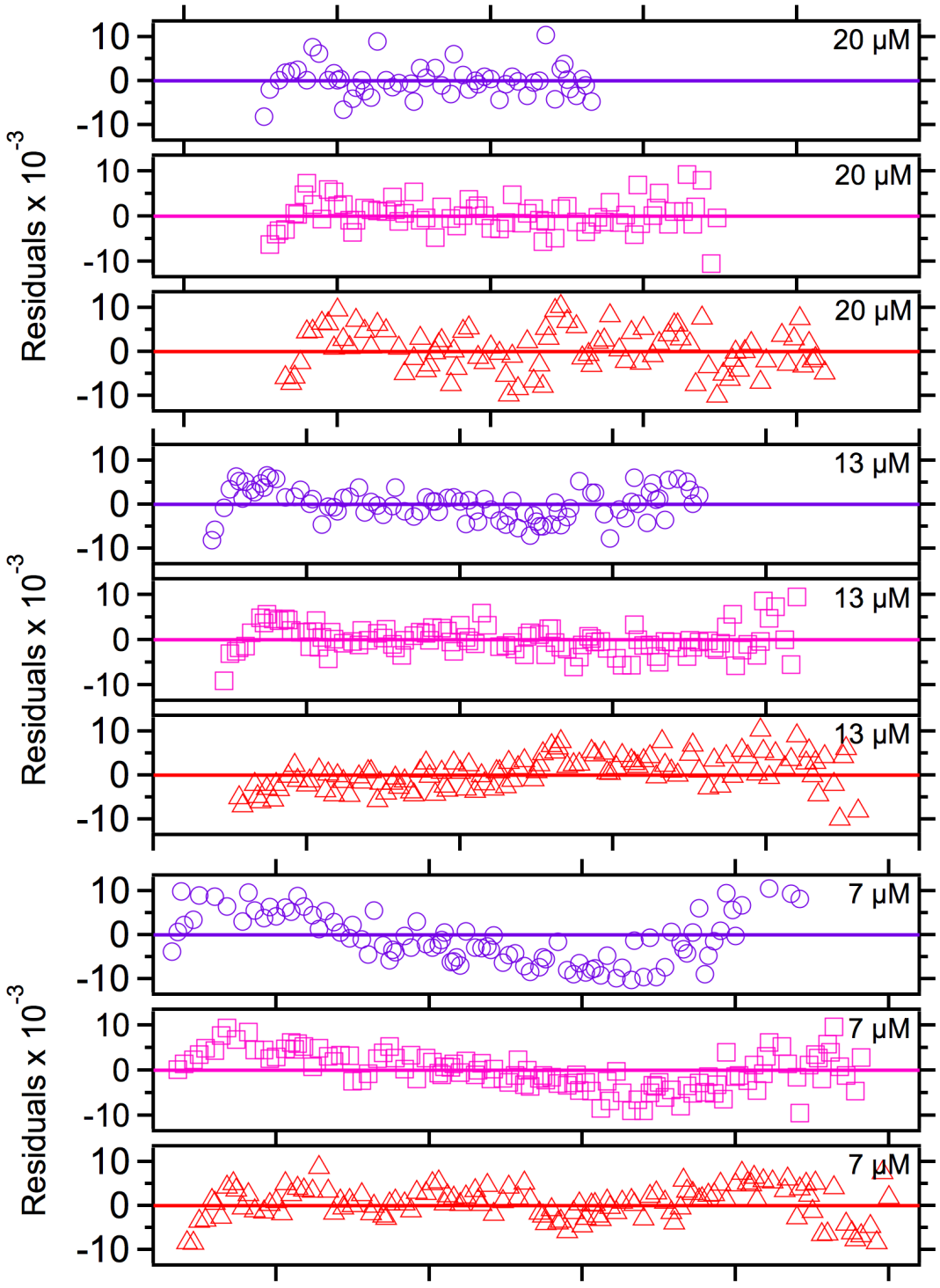


Figure 4.4B. Residuals for the fit to a monomer-trimer association scheme shown in Figure 4.4A. Symbols are as follows: purple circles, 20,000 rpm; magenta squares, 24,500 rpm; red triangles, 30,000 rpm.

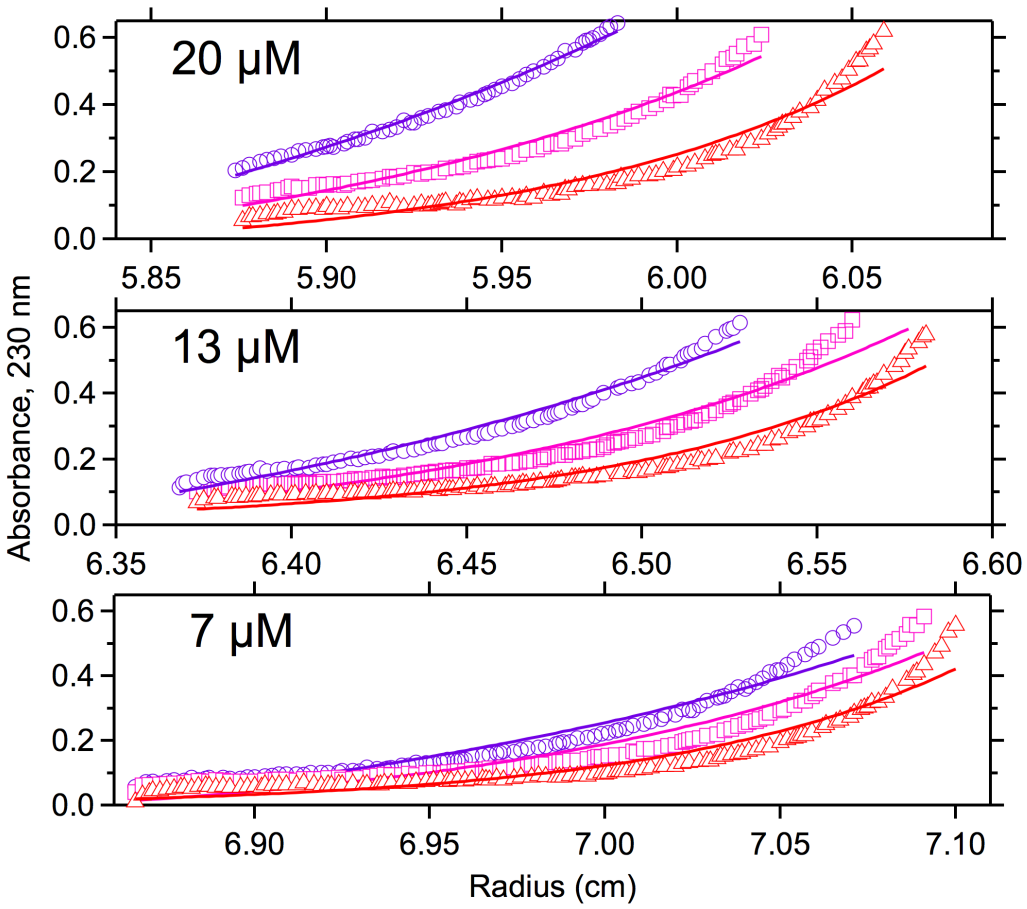


Figure 4.4C. Single species monomer does not fit the data. The same data as in figure 4.4A fit to an obligate monomer model cannot describe the SE data. Symbols are as follows: purple circles, 20,000 rpm; magenta squares, 24,500 rpm; red triangles, 30,000 rpm. The residuals for this fit are shown in figure 4.4D.

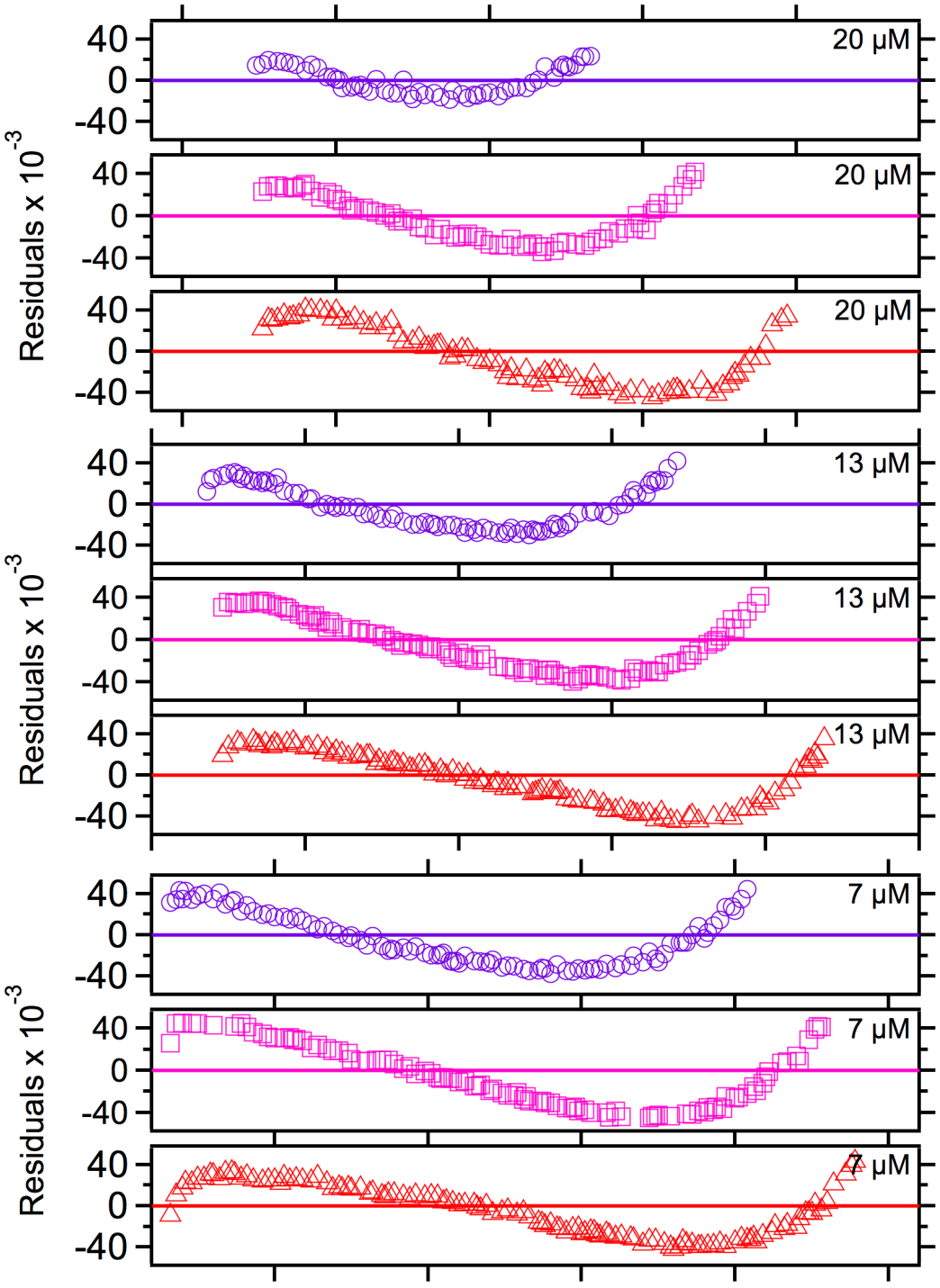


Figure 4.4D. Residuals for the obligate monomer model are systematic. Raw data and fits are shown in Figure 4.4C. Symbols are as follows: purple circles, 20,000 rpm; magenta squares, 24,500 rpm; red triangles, 30,000 rpm.

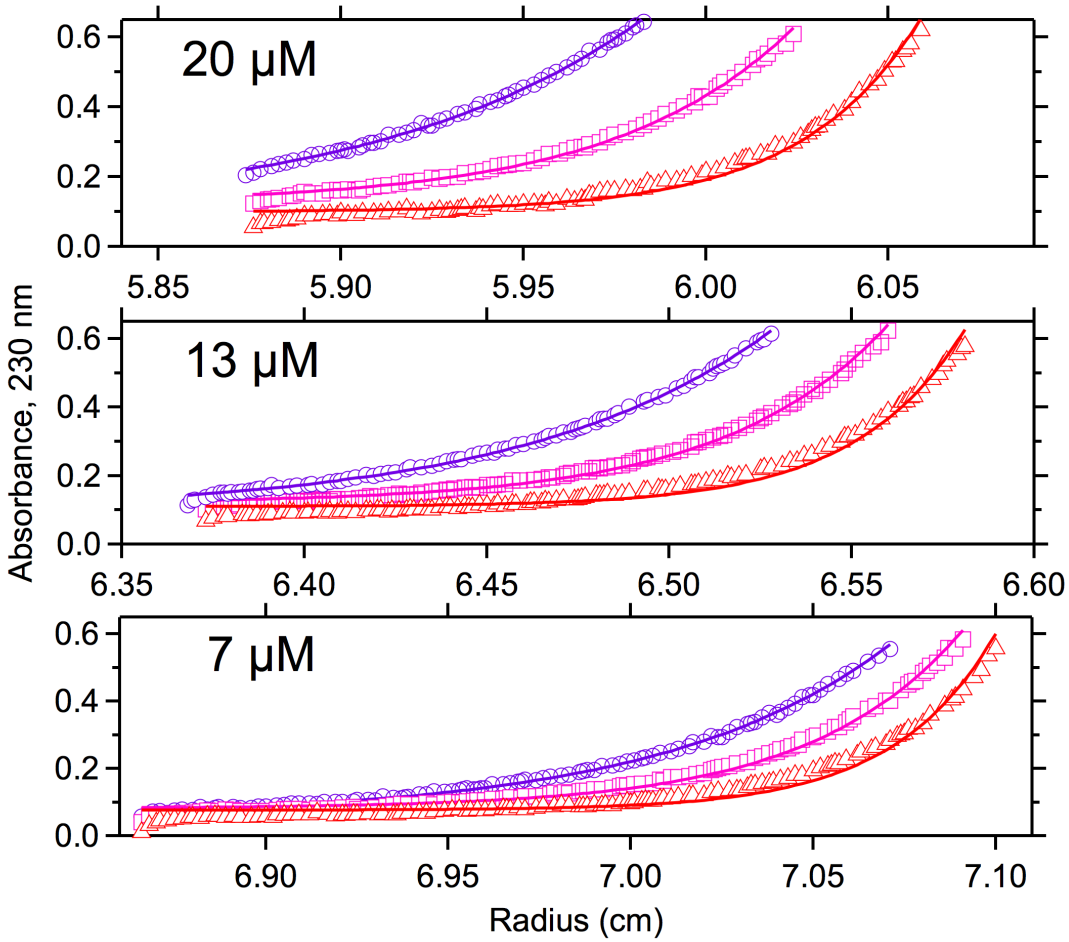


Figure 4.4E. Single species trimer does not fit the data The same data as in figure 4.4A but fit to an obligate trimer model cannot describe SE data for Skp at all rotor speeds. Symbols are as follows: purple circles, 20,000 rpm; magenta squares, 24,500 rpm; red triangles, 30,000 rpm. The residuals for this fit are shown in figure 4.4F.

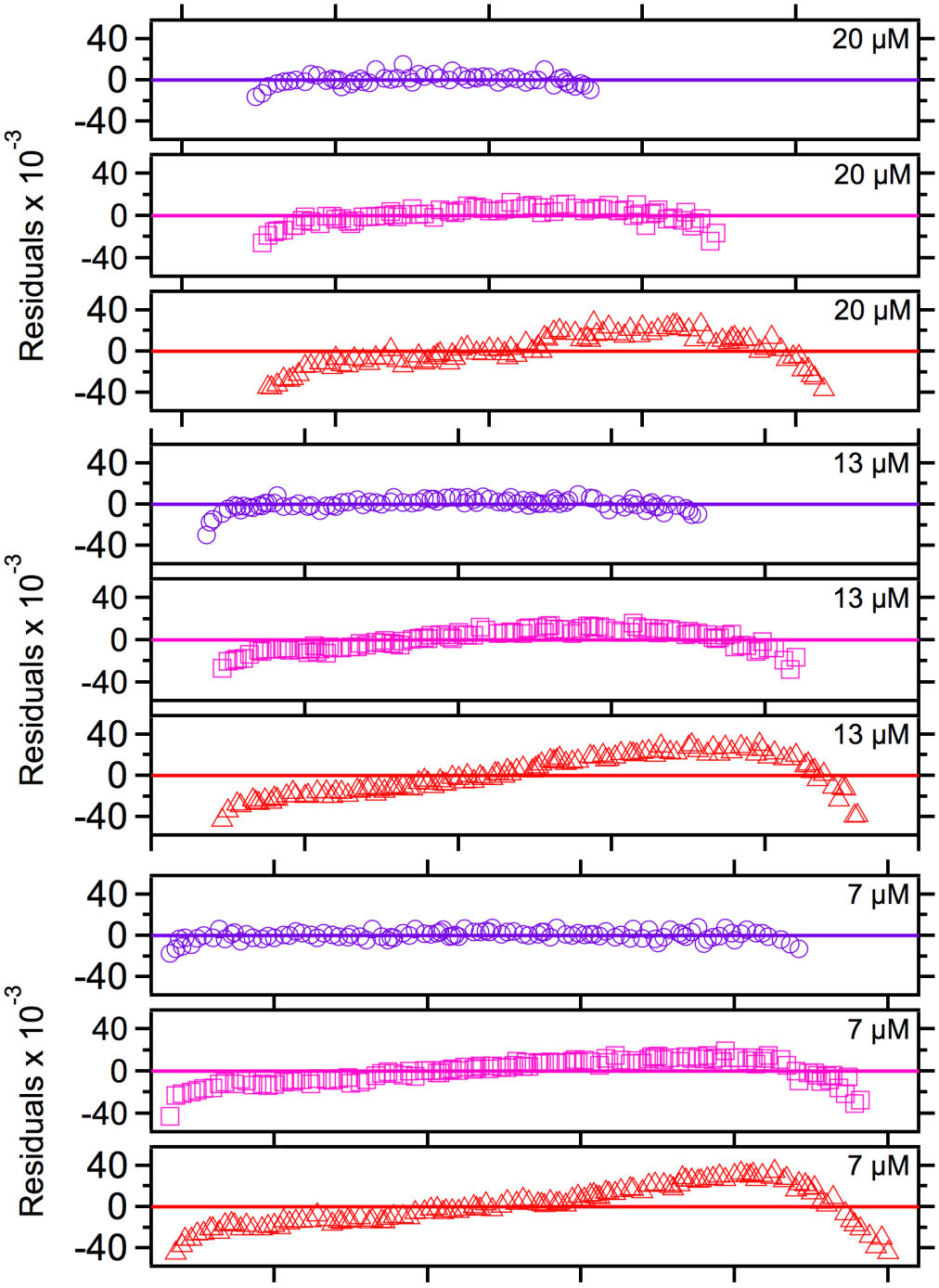


Figure 4.4F. Residuals for the single species trimer model are systematic at most concentrations and rotor speeds. Raw data and fits are shown in Figure 4.4E. Symbols are as follows: purple circles, 20,000 rpm; magenta squares, 24,500 rpm; red triangles, 30,000 rpm.

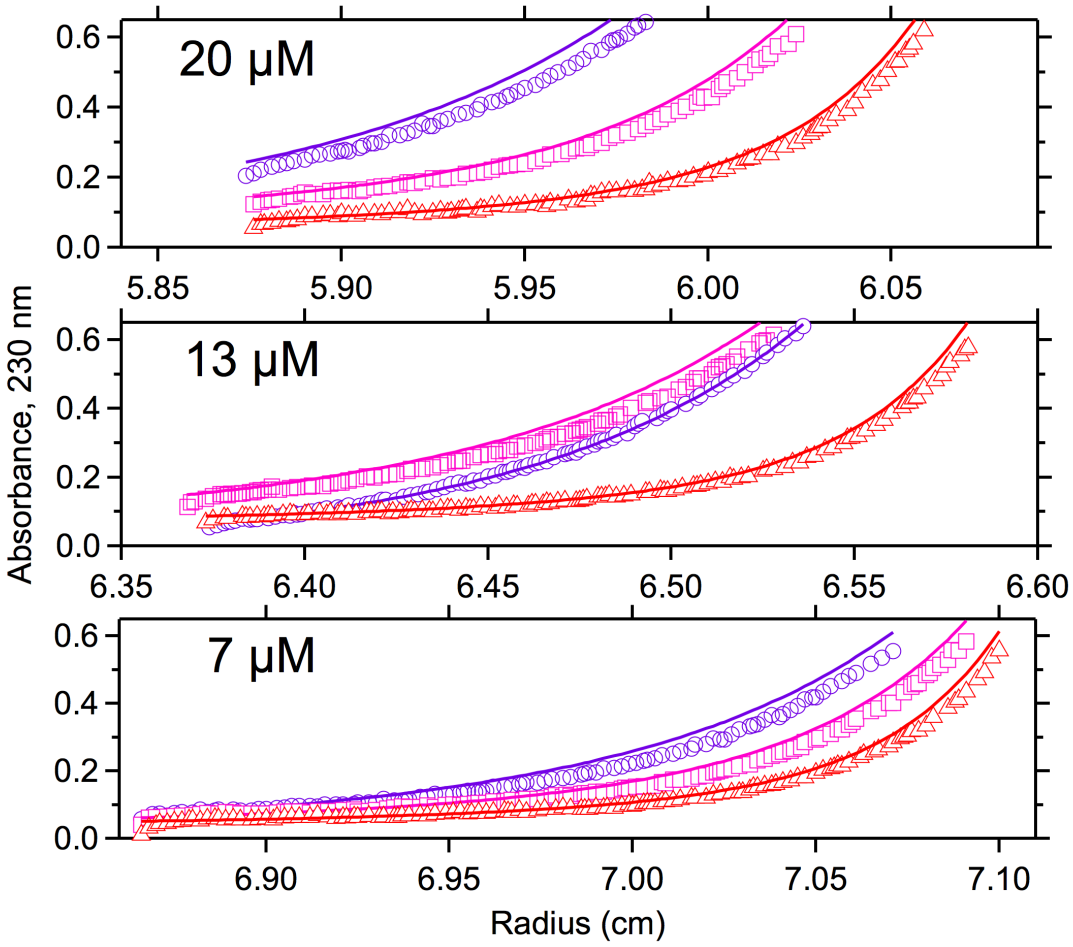


Figure 4.4G. A monomer-dimer-trimer association model is not an improvement over monomer-trimer. The same data as in figure 4.4A, fit to a monomer-dimer-trimer association scheme. Purple circles, 20,000 rpm. Symbols are as follows: purple circles, 20,000 rpm; magenta squares, 24,500 rpm; red triangles, 30,000 rpm. The residuals for this fit are shown in figure 4.4H.

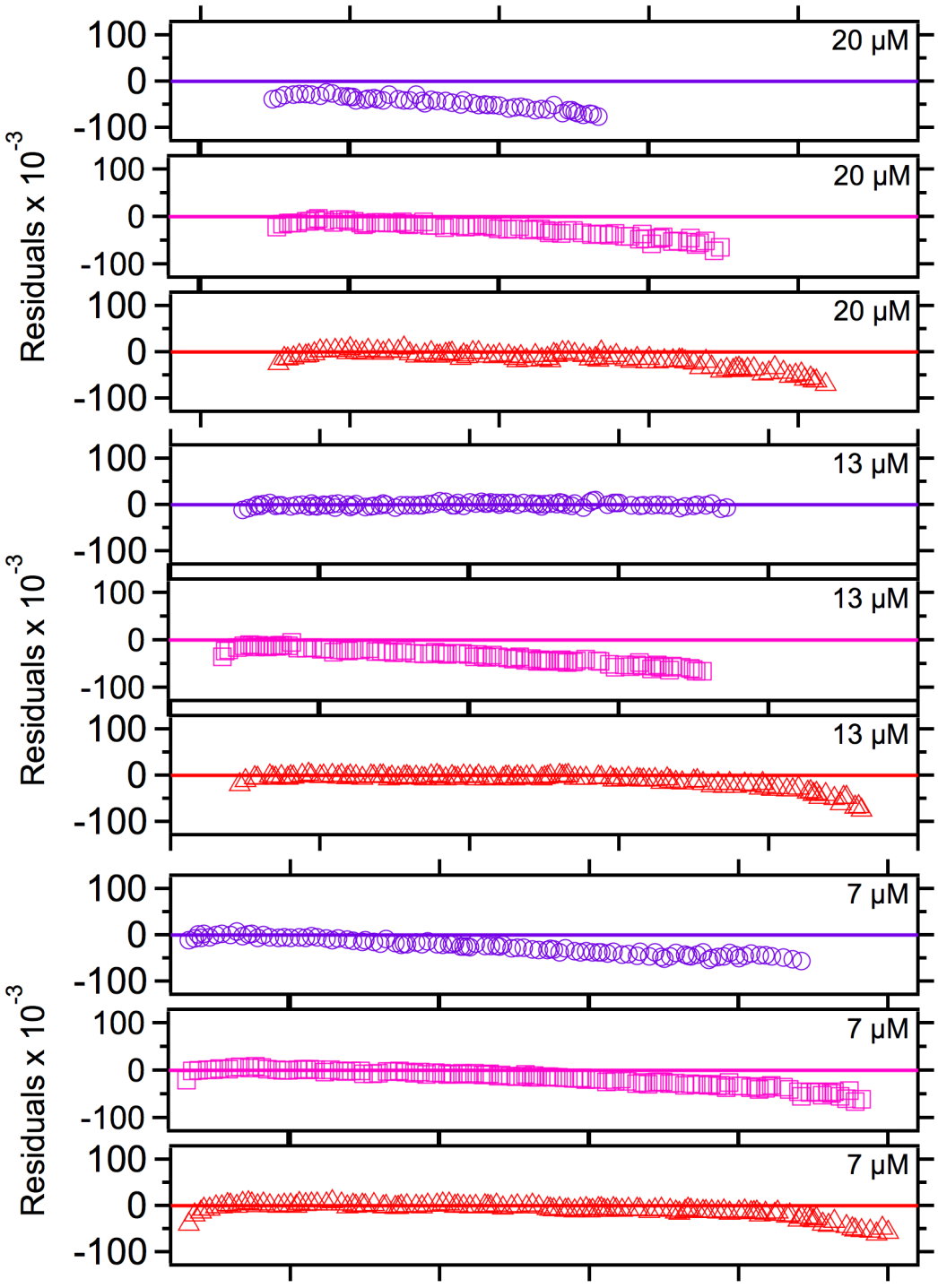


Figure 4.4H. Residuals for the fit shown in figure 4.4G are systematic for most Skp concentrations and rotor speeds. Symbols are as follows: purple circles, 20,000 rpm; magenta squares, 24,500 rpm; red triangles, 30,000 rpm.

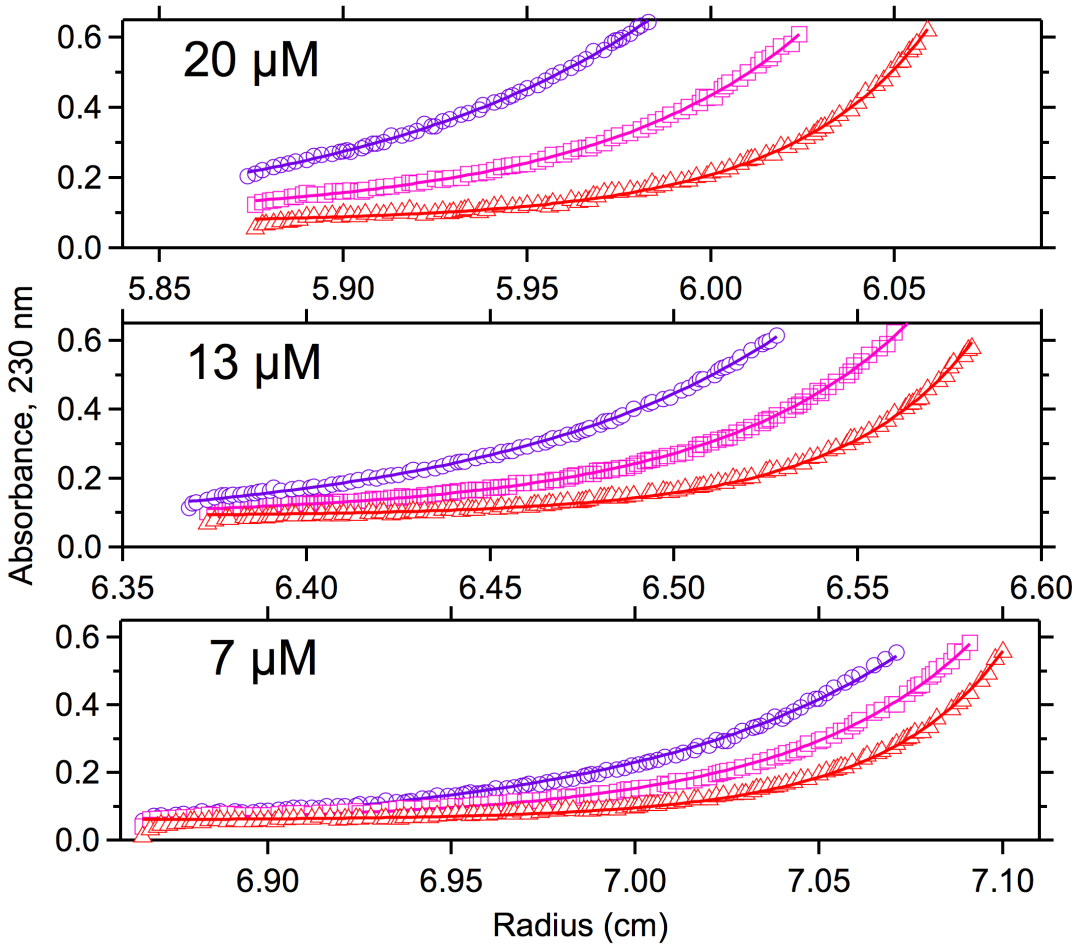


Figure 4.4I. A dimer-tetramer scheme does not improve the fit. The same data as in figure 4.4A fit to a dimer-tetramer association scheme. In addition, this association model is not compatible with the crystal structures of Skp. Symbols are as follows: purple circles, 20,000 rpm; magenta squares, 24,500 rpm; red triangles, 30,000 rpm. The residuals for this fit are shown in figure 4.4J.

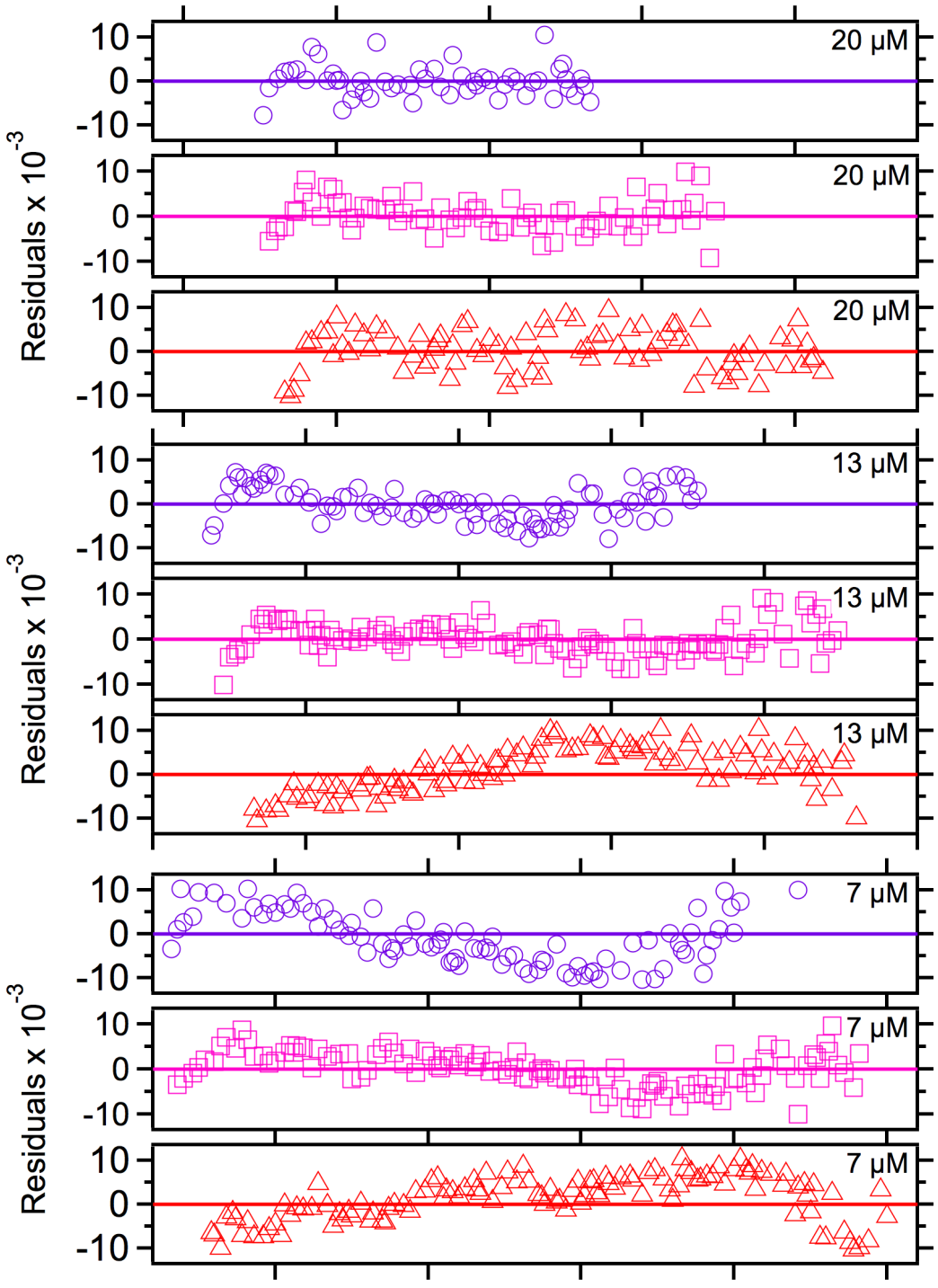


Figure 4.4J. Residuals for the fit shown in figure 4.4I. Symbols are as follows: purple circles, 20,000 rpm; magenta squares, 24,500 rpm; red triangles, 30,000 rpm.

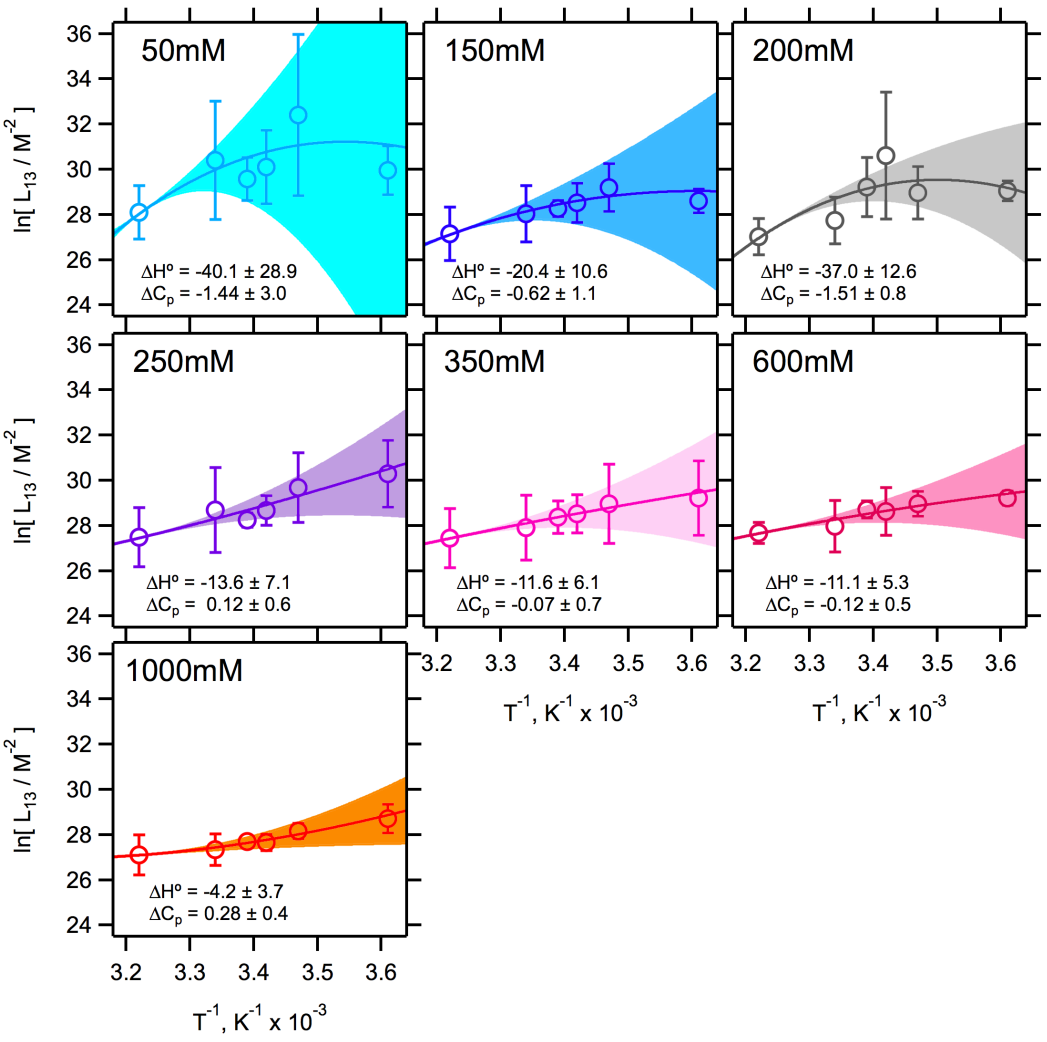


Figure 4.5A. The temperature dependent data for Skp trimerization fit to the integrated van't Hoff equation shows only a modest heat capacity change. The shaded regions indicate confidence boundaries for ΔC_p ± one standard deviation (SD) calculated with ΔH° fixed to the bootstrapped estimate. Concentrations in the upper-left-hand corner of each plot correspond to the [NaCl] in each condition. The bootstrap estimates for ΔH° and ΔC_p (both in units of kcal mol^{-1} K^{-1}) with their associated standard deviations are printed in the lower-left-hand corner of each plot.

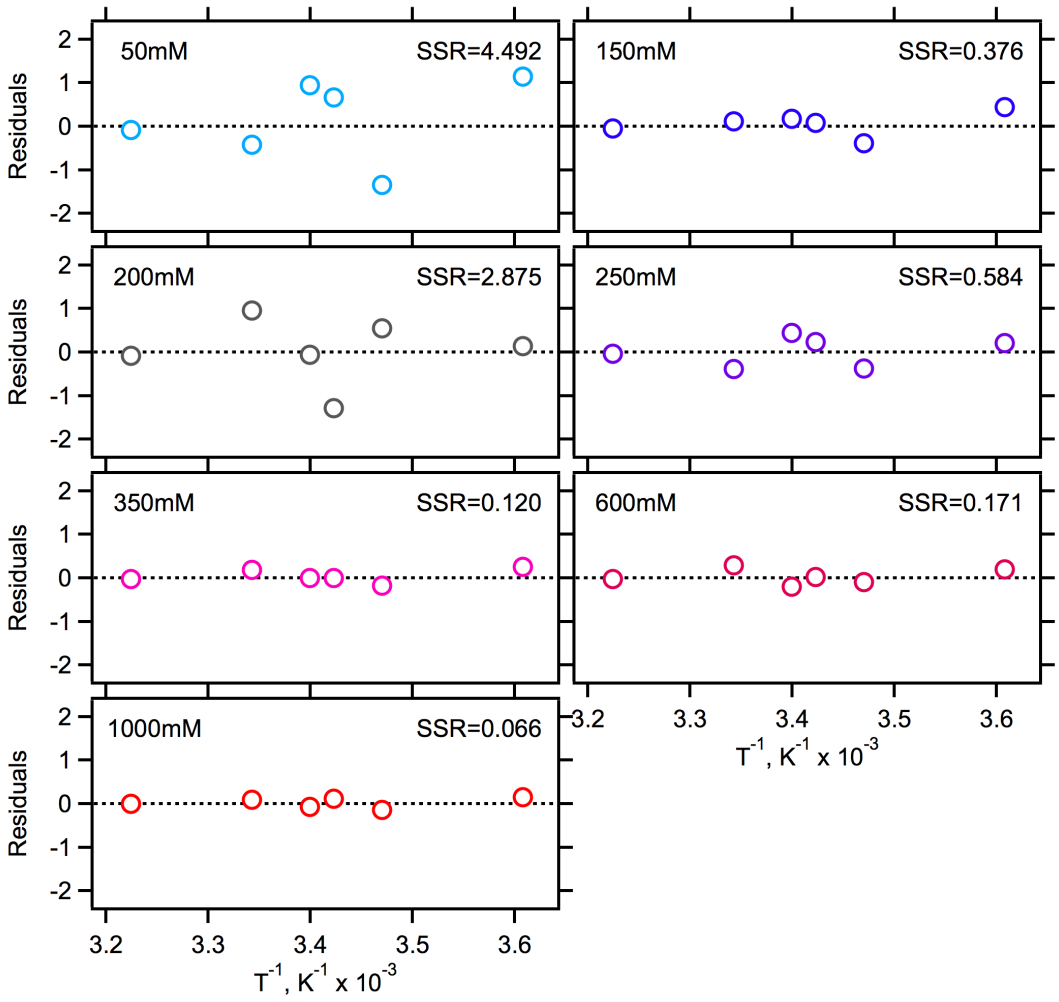


Figure 4.5B. Residuals and associated sums of squares of residuals (SSR) for temperature dependent data for Skp trimerization from the fits to the integrated van't Hoff equation. Concentrations are for [NaCl] in each condition as indicated.

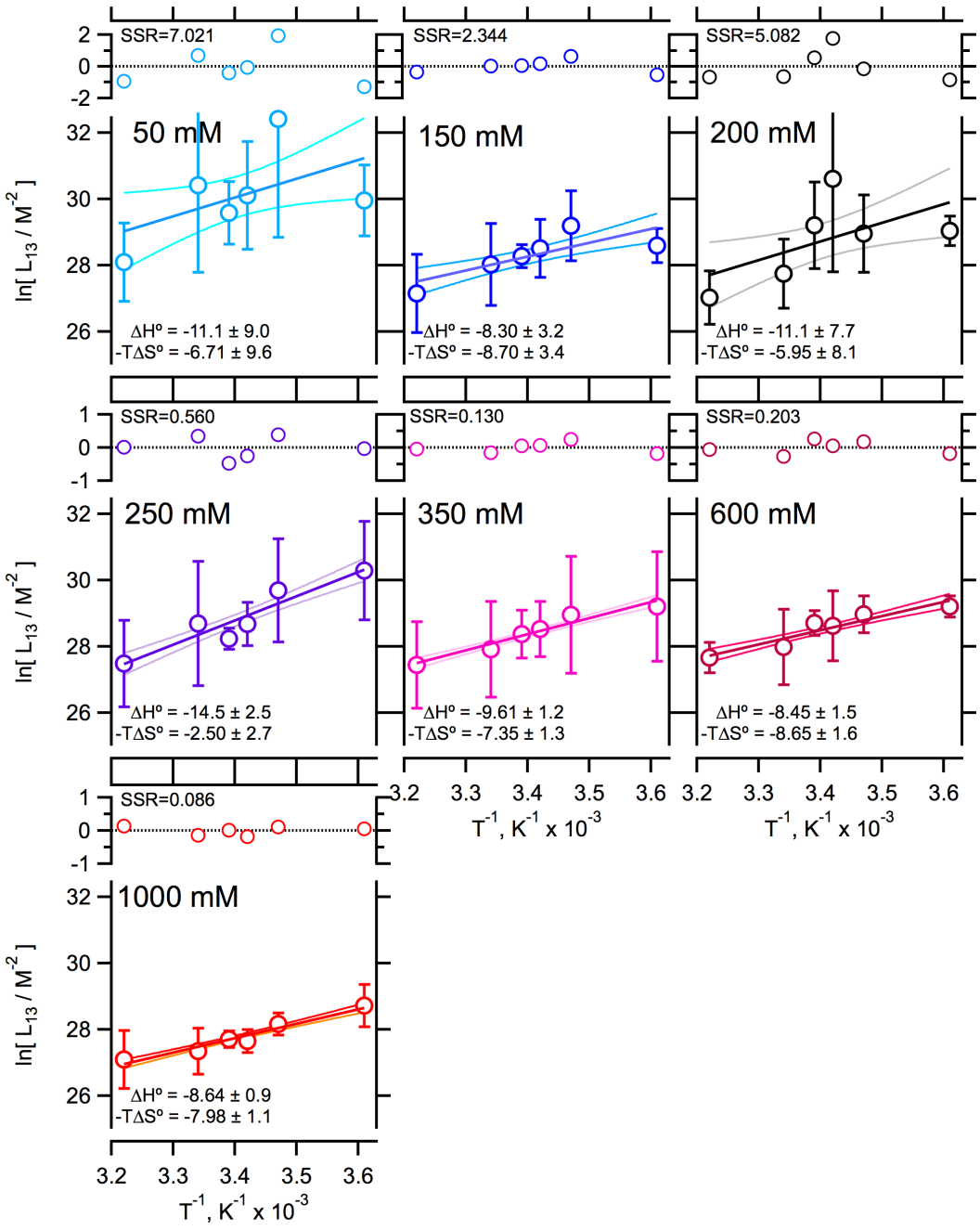


Figure 4.5C. The temperature-dependent data for Skp trimerization fit to the linearized van't Hoff equation. Confidence bands for the best fit of ln[L_{13}] ± one standard deviation. Concentrations indicated in each plot correspond to the [NaCl] in each condition. The estimates for ΔH° (units of kcal mol^{-1}) and -TΔS° (units of kcal mol^{-1}) with their associated SD are printed in the lower-left-hand-corner of each plot.

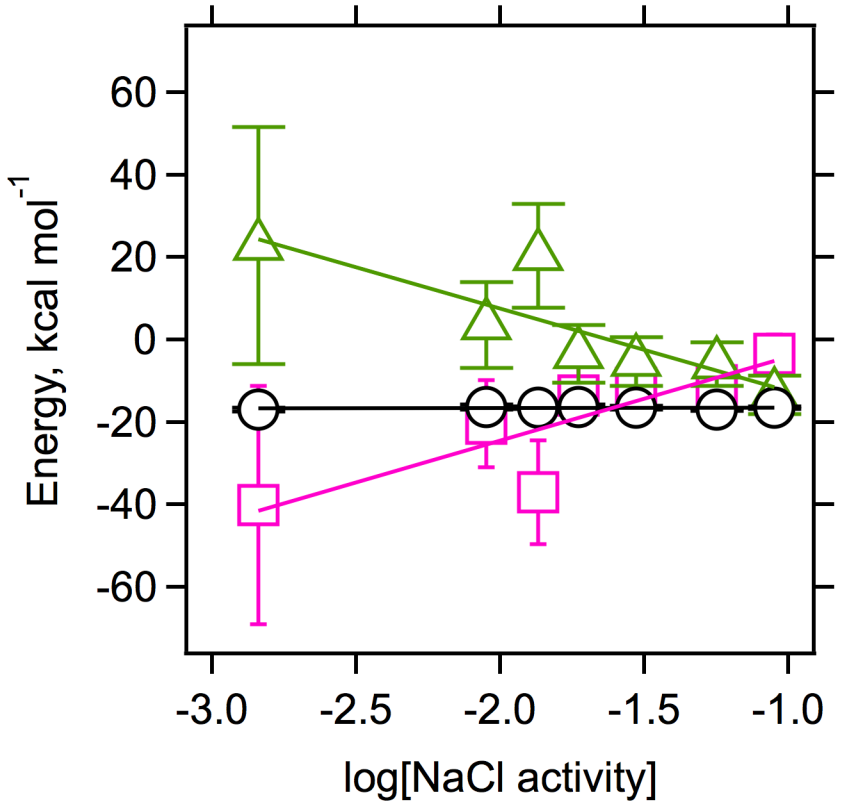


Figure 4.6. Thermodynamic data from integrated van't Hoff analyses for Skp trimerization at 37 °C, pH 8.0 calculated from seven different salt conditions indicates enthalpy-entropy compensation with increasing NaCl concentration. Symbols are as follows: magenta squares, $\Delta H°_{37\ °C}$ from the van't Hoff fits; green triangles, $-T\Delta S°$ at 37 °C calculated from the Gibbs equation as $-T\Delta S°_{37\ °C} = \Delta G°_{37\ °C} - \Delta H°_{37\ °C}$. Magenta and green lines are fits to guide the eye. Black circles, experimentally measured $\Delta G°_{37\ °C}$ (see figure 2A, purple).

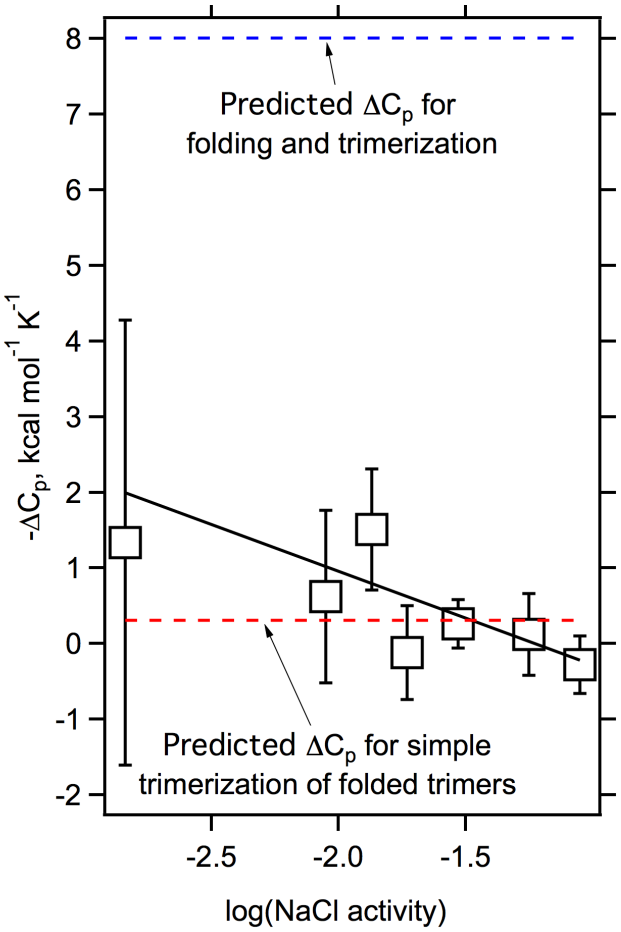


Figure 4.7. The ΔC_p estimates from the van't Hoff analysis are consistent with folded Skp monomers in equilibrium with the Skp trimer. The ΔC_p estimated at a reference temperature of 37 °C (as reported from the integrated van't Hoff fits in figure 4.5A) are plotted in squares as a function of NaCl concentration. The linear fit shows that all ΔC_p values are within error of the value ΔC_{13} = -0.31 kcal mol^{-1} K^{-1} calculated for trimerization of three well-folded Skp monomers (red dotted line). The blue dotted line indicates the expected value of ΔC_{total} = -8.01 kcal mol^{-1} K^{-1} calculated for the folding of three Skp protomers upon by trimerization (see section 4.2 and table 4.6).

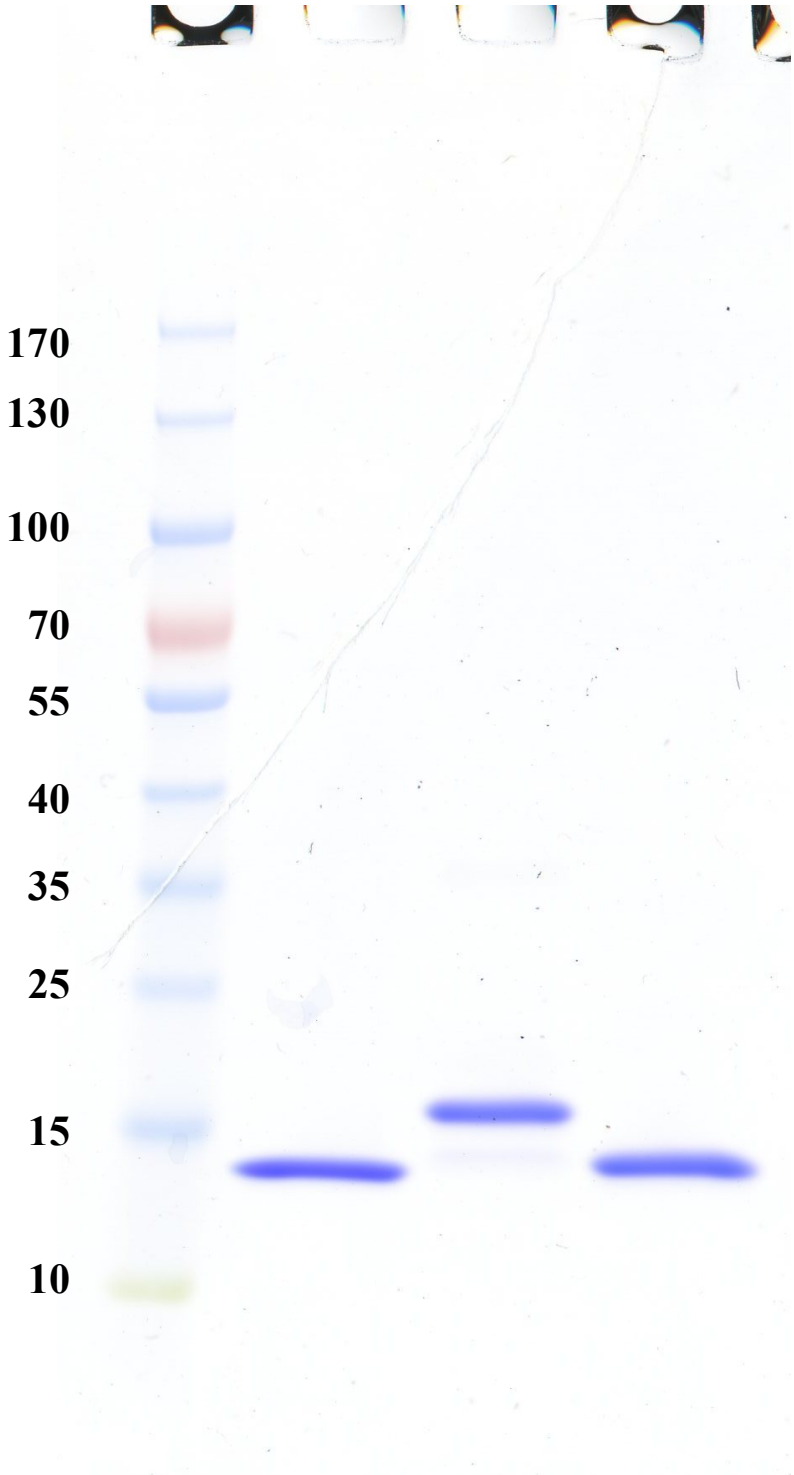


Figure 4.8. Representative SDS-PAGE gel displaying the high purity of the Skp protein samples used in this study. Molecular weight markers are marked with their associated molecular weight in kilodaltons. Lane 2, Skp-His (used for all measurements).

Lane 3, His-Skp. Lane 4, WT Skp, produced by digesting the protein His-Skp in lane 2 with TEV protease.

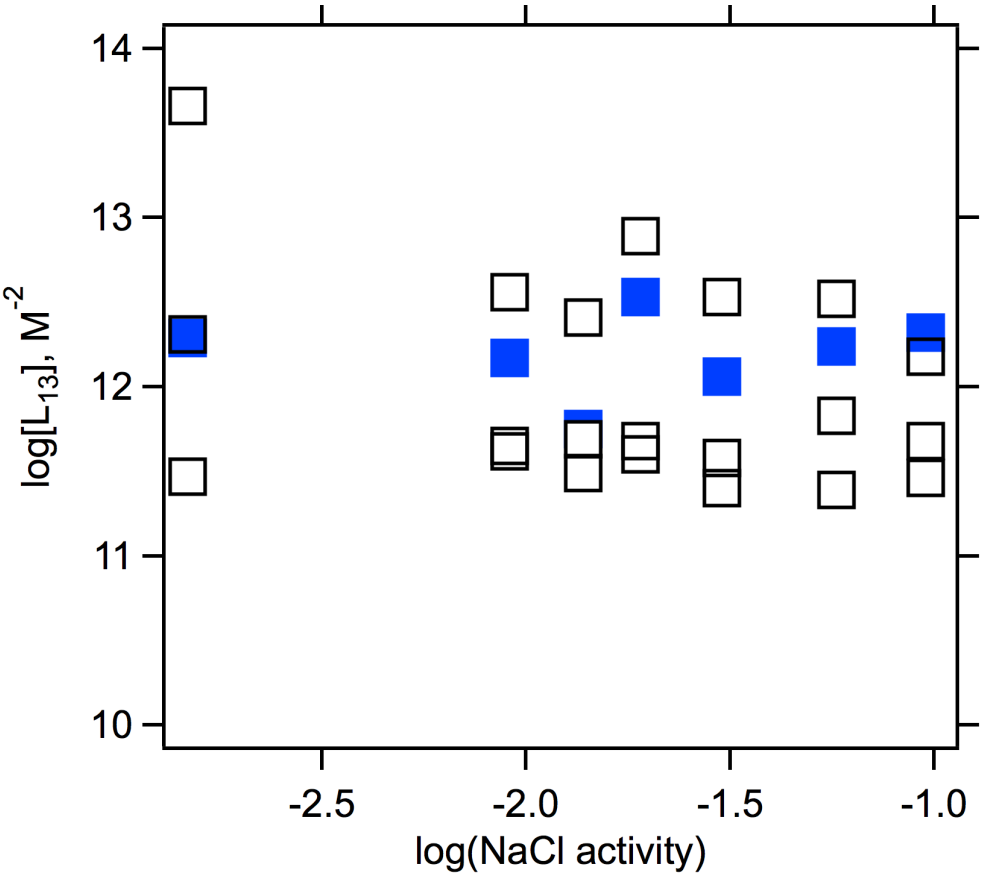


Figure 4.9. The C-terminal histidine tag has no effect on Skp-His trimerization energetics. Replicate data for Skp-His (open squares) shows comparable trimerization energetics to that of the control WT Skp lacking the 6-Histidine C-terminal tag (solid blue squares). Data are for 26 °C and pH 8.0 and all seven NaCl concentrations used in this study (left to right, 50, 150, 200, 250, 350, 600, and 1000 mM). The magenta data plotted in figure 2A show the mean and standard error of the three open squares shown here at each NaCl concentration.

Chapter 5
Skp is More Like a Buffer than a Chaperone for Outer Membrane Proteins

5.1 Introduction

Some proteins form a cage around a substrate, or client, that prevents their aggregation. A classic example is GroEL-GroES, which encloses clients in a barrel structure (Gupta, Haldar et al., 2014). In contrast, other proteins like trigger factor have antiaggregation activity (AA) but do not form cages (Saio, Guan et al., 2014). AA (section 1.2 and figure 1.5) is kinetic competition between client binding and aggregation. AA differs from chaperone activity (CA), which is acceleration of client folding kinetics. Together, AA and CA are called holding activity (Mattoo and Goloubinoff, 2014).

The Seventeen Kilodalton Protein (Skp) forms trimers in the periplasm of *Escherichia coli* that are believed to cage unfolded outer membrane proteins (uOMP) to provide holding activity (Burmann, Wang et al., 2013). However, Skp exists in both trimeric and monomeric forms. The trimeric form could increase if Skp is upregulated by uOMPs (Sandlin, Zaccai et al., 2015). This suggests a "switching hypothesis", which predicts that Skp monomers are inactive, while Skp trimers provide either AA or holding activity.

Even if switching occurs, the size of a uOMP contained within Skp may limit Skp function. For example, a uOMP may fail to fit inside the Skp trimer, leading to aggregation (figure 1.7) (Walton, Sandoval et al., 2009). Furthermore, Skp and each uOMP client adapt to each other structurally (Zaccai, Sandlin et al., 2016). Such

structural adaptation could also imply a variation in AA. Variation of AA with client type was shown for antibody fragments (Entzminger, Chang et al., 2012). For these fragments, Skp AA also correlated with CA, implying that Skp is a holding chaperone (Entzminger, Chang et al., 2012). Still, nonnative antibodies may behave differently from hydrophobic uOMPs. These observations led me to ask how Skp AA varies with uOMP structure, and whether Skp is truly a holding chaperone for uOMPs.

To test the switching hypothesis, I employed sedimentation velocity (SV) to assay the AA of a monomeric Skp construct, dubbed FAI (Full Alanine Interface). In support of switching, I found that FAI AA is limited compared to wild-type Skp. Next, to check whether Skp is a holding chaperone for uOMPs, I measured Skp AA and CA for uOmpA, uOmpLa, and uBamA-TM. Surprisingly, Skp AA varied significantly between these uOMPs, but this AA did not correlate with CA measured by SDS-PAGE. Finally, I show that uOmpLa and uBamA aggregate above the ~nM concentration that Skp binds. Based on these results, I suggest a new model: Skp primarily functions in *E. coli* to keep free uOMP concentrations low, while it possesses only minor function as a chaperone.

5.2 Materials and Methods

Antiaggregation Activity (AA) Defined in Terms of Observables in Sedimentation Velocity (SV)

To present AA data in terms of the observables in SV, I first define AA as a ratio

$$\theta(\mathrm{t}) \equiv \frac{[M]_{-}(\mathrm{t})-[M]_{-}(0)}{[M]_{+}(\mathrm{t})-[M]_{+}(0)} \qquad \text{Equation 5.1}$$

In equation 5.1, $[M]_{+}$ is the concentration of monomeric uOMP when Skp is present, and $[M]_{-}$ is the monomeric uOMP when Skp is absent. The variable t indicates

time, and M(0) denotes values at time zero. This choice for an AA parameter makes sense for the following three reasons. First, $\theta(t)$ is usually positive at some time t>0, because both the numerator and denominator of $\theta(t)$ are negative (i.e., loss of monomer). Second, $\theta(t)$ increases with the speed of uOMP aggregation (i.e., as $[M]_-(t) - [M]_-(0)$ increases in magnitude) and increases the more that Skp slows such aggregation (i.e., as $[M]_+(t) - [M]_+(0)$ increases in magnitude). Third, it is a ratio of apparent rates, because

$$\frac{[M]_-(t)-[M]_-(0)}{[M]_+(t)-[M]_+(0)} = \frac{\Delta[M]_-}{\Delta[M]_+} = \frac{\Delta[M]_-/\Delta t}{\Delta[M]_+/\Delta t} = \frac{\text{"Rate without Skp"}}{\text{"Rate with Skp"}} \qquad \text{Equation 5.2}$$

To estimate $[M]_+(t)$ and $[M]_-(t)$ from SV data, I integrate over the absorbance distribution, g(S*), of the late-sedimenting species. To perform this integration (equation 5.3), I define s as a dummy variable for S* in g(S*) and t_1 as the time at which the centrifuge run is started. In equation 5.3, s_1 and s_2 represent the upper and lower limits of integration:

$$C_{+/-}(t_1) = \int_{s_1}^{s_2} g(s, t_1)ds \qquad \text{Equation 5.3}$$

$C_{+/-}(t_1)$ is the estimate for the absorbance contribution of the monomeric uOMP, assuming the reaction is frozen at the start of the run. To estimate $M_+(0)$ and $M_-(0)$ for equation 5.1, I define C(0) as the total absorbance contribution from the uOMP at the time of the first scan. Next, I take the following fraction (equation 5.3),

$$\alpha(t_1) = [C(t_1) - C(0)]_{\text{uOMP only}}/[C(t_1) - C(0)]_{\text{uOMP+Skp}} \qquad \text{Equation 5.3}$$

At 280 nm, Skp has insignificant absorbance relative to the uOMP, giving an approximation of uOMP monomer concentration without and with Skp at time t_1 (equation 5.4)

$$\alpha(t_1) \approx \frac{\Delta A^-_{280,\ \text{late-sedimenting uOMP}}(t_1)}{\Delta A^+_{280,\ \text{late-sedimenting uOMP}}(t_1)} \qquad \text{Equation 5.4}$$

Beer's law can be used to show that equation 5.4 is a reasonable estimate for θ (equation 5.1).

The Extinction Coefficient of SAI (and FAI) at 230 nm Determined by Refractive Index

Concentrations of Skp from 0-100 μM were estimated from the theoretical extinction coefficient for Skp at 280 nm of 1490 M^{-1} (Sandlin, Zaccai et al., 2015). These samples of Skp were measured for their refractive index (RI) using a Reichert AR7 Series Automatic Refractometer (Reichert). The resulting standard curve was used to determine the dependence of protein RI on concentration (figure 5.1A). Using this dependence, I used another standard curve to estimate the extinction coefficient of SAI at 230 nm (figure 5.1B). The extinction coeffient of SAI was estimated at ε_{230} = 30400 ± 2800 M^{-1} cm^{-1}. Owing to limited sample availability at the time of writing, the same extinction coefficient was also used for FAI.

Sedimentation Equilibrium of Skp-FAI and Skp-SAI

Methods followed section 4.2, except that the molecular weights used were 16553.8 for SAI and 16329.6 for FAI, and the respective degrees of freedom were 764 and 566.

Circular Dichroism (CD) Measurements of Skp, Skp-FAI and Skp-SAI

CD measurements were performed as described (Danoff and Fleming), except that the scan range was from 205 to 280 nm.

Sedimentation Velocity Assay for Antiaggregation Activity (AA)

Concentration of uOMPs was determined from the theoretical extinction coefficients (Moon, Zaccai et al., 2013). Protein samples stored at -80 °C and 60-80 μM uOMP in 20 mM Tris-HCl and 6 M urea, pH 8.0 (Sigma) in 1.5 mL microfuge tubes were thawed for 15 minutes, gently resuspended without vortexing, and allowed to incubate at ambient temperature for 1 hour. If necessary, stock concentrations were adjusted to 50 μM. Both uOmpA and uBamA-TM samples contained 1-2 mM TCEP at all steps, and 2 mM TCEP in the AA assay.

To begin the reactions, uOMP samples were rapidly diluted 25-fold with 20 mM Tris–HCl, 156.25 mM NaCl, and 62.5 mM urea, with or without 18.75 μM Skp. Immediately on dilution I repipetted six times while swirling with the pipette tip, then gently inverted the tube four times. Final concentrations for an assay containing Skp and TCEP were thus 2 μM OMP, 18 μM Skp, 20 mM Tris-HCl, 150 mM NaCl, 2 mM TCEP-HCl and 300 mM urea at pH 8.0. Based on estimates of Skp solubility as a function of urea (see Chapter 3), I estimate that 89% of Skp was functionally active at 300 mM urea.

To assay AA, this reaction was allowed to equilibrate at ambient temperature for either 30 or 40 minutes (figure 5.9) while being loaded into a 2-sector centerpiece. Loading was performed at ambient temperature (20-24 °C), and the centrifuge set to vacuum within 15 minutes of the start of the reactions. At 30 or 40 minutes post-dilution (mpd), centrifugation was performed at 201,600 g at 20 °C or 25 °C (table 5.2) with an An-60Ti rotor in a Beckman XL-A Analytical Ultracentrifuge (AUC). Data were collected as OD280, where uOMPs are detectable and Skp contributed an absorbance of

less than 0.03. Typically, 400+ scans were collected with a step size of 0.05 for AA assays, or 0.03 for determination of hydrodynamic radius (R_H).

Using the resulting scans, the late-sedimenting species were identified using DCDT+ software by Philo (2006). The late-sedimenting species were fit to a single Gaussian function, with $S_{20,W}$* fixed to the experimentally extrapolated value (figure 5.11) at 0.300 M urea and molecular weights fixed to their theoretical values (see table 5.2). Buffer parameters were determined as described (Sandlin, Zaccai et al., 2015). The absorbance contribution of the monomeric species was estimated from the resulting value of C0 from fitting. To estimate the total absorbance of samples, I took the average of three data points near 6.5 cm from the first scan.

Sedimentation Velocity Determination of $S_{20,W}$ for uOmpLa and uBamA-TM*

OMP stocks stored at -80 °C were handled as described above for the AA experiments. 60 μM OMP stock in 20 mM Tris-HCl and 4.25-7.5 M urea at pH 8.0 was diluted with a similar buffer to <10 μM to achieve OD280 >0.4. At 30 mpd, SV was performed as described above. DCDT+ was used to estimate $S_{20,W}$* for the late-sedimenting species, as clearly identified from the 6.0 or 7.5 M urea condition, where one symmetrical, Gaussian g(S*) curve was apparent throughout the runs.

Design, Expression, and Purification of Skp and uOMP Variants

Skp was expressed and purified as described (Sandlin, Zaccai et al., 2015). Constructs SAI, FAI, and OmpLa-WF were submitted to GeneWiz, Inc. for synthesis and cloning and subjected to GeneWiz quality-checking protocols. Skp variants SAI and FAI

were cloned into the same pOPINe vector as Skp, and expressed and purified as previously described for the wild-type variant (Sandlin, Zaccai et al., 2015). OmpLa-WF was cloned into a pET11a expression vector, and this vector was used to transform HMS-174 cells.

Expression and purification of OmpLa variants, OmpA, and BamA were performed as described previously (Burgess, Dao et al., 2008), except for the following changes. 0.2 mg/mL DNAse (Roche) was added to the first wash step, and the resuspended pellet was allowed to incubate with rocking for 30 minutes at ambient temperature before continuing with the second wash. After solubilization of inclusion body pellets in 20 mM Tris-HCl, 8 M urea (pH 8.0), supernatants were clarified by centrifugation in a JA-21 rotor in a Beckman centrifuge at 46,641 g to remove debris. The resulting supernatant was filtered with a 0.22 μm filter. If necessary, samples were concentrated with 15 mL Amicon filter concentrator (Millipore) with a 10 kDa cutoff according to the manufacturer's instructions, to >60 μM.

The BamA-TM construct was also expressed with a pET11a vector in HMS-174 cells. However, upon application of our usual protocols for uOMP inclusion body purification (see above) I observed that most BamA-TM was found in the soluble fraction after addition of lysis buffer (figure 5.2C). In both size-exclusion chromatography and SV, the highly pure BamA-TM did not form high-molecular weight aggregates as expected, but formed what appeared to be soluble low-molecular weight aggregates. Concerned that the sample was contaminated with lipid, I purified BamA-TM differently. After solubilization in lysis buffer, I subjected BamA-TM to methanol-chloroform extraction. The resulting white solid was re-solubilized in 20 mM Tris-HCl, 6 M Urea, 1

mM TCEP, pH 8.0, before clarification and concentration as described above. The resulting highly pure BamA-TM behaved in SV like an insoluble membrane protein, showing that contamination during preparation can drastically change the aggregation behavior of uOMPs in experiments. Figure 5.2 compares SV data for soluble Skp in GF, which sediments as monomer and trimer (figure 5.2A) (Sandlin, Zaccai et al. 2015), with BamA-TM (figure 5.2B) before and after chloroform-methanol extraction.

Fluorescence Binding Experiments

Methods followed those detailed in (Moon, Zaccai et al., 2013). Briefly, a solution containing 40 nM uOmpLa in urea and varying concentrations of FAI (figure 5.8) was placed in an ISS fluorometer with a Quantum Northwest temperature control unit set to 25 °C and a stir setting of 3. Data collection was performed in slow kinetics mode at an excitation wavelength of 295 nm and an emission wavelength of 341 nm. Data were fit to a single exponential to recover the steady-state fluorescence intensities. The resulting data were fit to a single-site binding model as described (Moon, Zaccai et al., 2013) to recover the apparent K_D.

Fluorescence Kinetics of uOMP Aggregation

uOMP aggregation assays were performed using an ISS fluorometer using the same settings as for the fluorescence binding experiments. To further reduce photobleaching, the excitation polarizer was inserted and set to 0 degrees, the lamp power was set to 14 A, and the shutter was closed between time points. The step-size of data collection was 5 seconds. Samples of uOMP (see above) at 50 μM in 6 M urea at pH 8.0

were rapidly diluted with a solution containing NaCl and varying concentrations of urea to achieve final uOMP concentrations ranging from 83 nM to 910 nM. For kinetics collected over 24-hours (figure 5.12E), the fluorometer cuvette was also covered with parafilm.

uOmpLa Aggregation Kinetics by Sedimentation Velocity

The methods used to perform kinetics measurements of uOmpLa aggregation (figures 5.12D and E) were similar to those used for the AA assays. However, varying concentrations of uOmpLa and urea were combined to achieve the final concentrations described in the caption of figure 5.12. To quantify the fraction of monomeric uOmpLa, DCDT+ was employed as in the AA assays, and the resulting estimate for C0 was divided by the total absorbance. This total absorbance was estimated by averaging the three points closest to 6.5 cm in the first velocity scan.

Development of a Hydrophobic Sequence Segregation Parameter, ξ

I sought to develop an analogue of kappa (κ) in terms of hydrophobicity. κ is related to the distribution of charge along a polypeptide chain. To start, I examined the form of κ. The following is taken from the derivation of κ (Das and Pappu, 2013). If N is the number of residues in a polypeptide, and N_+ is the number of cationic residues, the fraction positive is

$$f_+ = \frac{N_+}{N} \qquad \text{Equation 5.5}$$

Similarly, f- is the fraction of negative charges. The charge asymmetry is defined as

$$\sigma \equiv \frac{(f_+ - f_-)^2}{f_+ + f_-}$$ Equation 5.6

Next, define g as the length of a sliding window, called a blob, and let N_{blob} be the total number of blobs, such that

$$N_{blob} = N - g + 1$$ Equation 5.7

The squared deviations of the charge asymmetry of the blobs, relative to the overall charge asymmetry of the sequence is defined as

$$\delta \equiv \frac{\sum_{i=1}^{N_{blob}} (\sigma_i - \sigma)^2}{N_{blob}}$$ Equation 5.8

Also, let δ_{max} be the maximum possible value that δ can take for any arrangement of a polypeptide chain of a specified composition. With these assumptions, κ is defined as (Das and Pappu, 2013)

$$\kappa \equiv \left[\frac{\delta(g=5)}{\delta_{max}(g=5)} + \frac{\delta(g=6)}{\delta_{max}(g=6)}\right]/2$$ Equation 5.9

I envisioned that the new hydrophobic segregation parameter, dubbed ξ, could be based on an analogue of the σ used in equations 5.6 and 5.8. I named this hydrophobicity analog σ_H. However, the terms f in equation 5.6 are counts of charges, and hydropathy is scalar data. Also, hydropathy does not take on a positive and negative form as charges do.

Therefore, for the numerator of σ_H, I define a "solvophobicity per residue" term, H_{sol}. H_{sol} relates the relative hydropathy of a residue, as specified by the Moon-Fleming scale (Moon, 2011), to the thermal energy. In H_{sol}, hydropathy favors collapse, whereas thermal energy favors expansion. I define $\Delta\Delta G_{wl,Moon-shifted}°$ as the water-lipid transfer free energy of an amino acid, shifted relative to lysine, minus 0.01 kcal/mol. This last adjustment keeps the denominator of σ_H from becoming zero. Moreover, I define H_{sol}:

$$H_{sol} \equiv \begin{cases} \left|\Delta\Delta G_{wl,Moon,rel.\ K-0.01}{}^{\circ}\right|, & \text{if } \left|\Delta\Delta G_{wl,Moon-shifted}{}^{\circ}\right| > 4.61RT \\ 0, & \text{if } \left|\Delta\Delta G_{wl,Moon-shifted}{}^{\circ}\right| \leq 4.61RT \end{cases} \quad \text{Equation 5.10}$$

This choice selects residues that have a 100-fold higher tendency than lysine, relative to RT, to become buried.

Now I describe the choice for a denominator for σ_H. In equation 5.6, the denominator represents the total charge density along the chain. Charge density is the same as the "average charged". Therefore, the average hydrophobicity is the best choice of a hydrophobic analogue for charge density. Therefore, I define

$$H \equiv \left|\Delta\Delta G_{wl,Moon-shifted}{}^{\circ}\right| \quad \text{Equation 5.11}$$

as the hydrophobicity of each residue. Finally, define σ_H as

$$\sigma_H \equiv \frac{\langle H_{sol}\rangle^2}{\langle H\rangle} \quad \text{Equation 5.12}$$

were the brackets indicate the average over all residues. The rest of the terms in my ξ parameter are completely analogous to those above for calculating κ, such that

$$\delta_H \equiv \frac{\sum_{i=1}^{N_{blob}}(\sigma_{Hi}-\sigma_H)^2}{N_{blob}} \quad \text{Equation 5.27}$$

and

$$\xi \equiv \left[\frac{\delta_H(g=5)}{\delta_{H,max}(g=5)} + \frac{\delta_H(g=6)}{\delta_{H,max}(g=6)}\right]/2 \quad \text{Equation 5.28}$$

The Python code I used for calculation of ξ is provided in appendix II. Table 5.1 lists the resulting ξ values for all the uOMPs commonly studied by the Karen Fleming lab.

Delayed-Folding Experiment to Assay Skp Chaperone Activity

To estimate CA under the same conditions as the SV AA assays, I performed delayed folding experiments similar to those used in chapter 3 (figure 3.1). I started with

the same conditions as the AA assays. However, instead of using SV at 30 mpd, I added the aggregating uOMPs to diC_{11}PC large unilamellar vesicles (LUVs). Final folding conditions were 1 mM LUVs (556:1 lipid:protein ratio), 1.8 μM uOMP, 16.2 μM Skp, 20 mM Tris-HCl, 150 mM NaCl, and 300 mM urea (pH 8.0). All folding reactions contained 4 mM TCEP, except for the reaction containing OmpLA.

5.3 Results

Mutation of the Trimerization interface of Skp Reduces Trimer Formation

The switching hypothesis predicts that trimerization is necessary for AA. Therefore, I sought to test AA for a constitutively monomeric form of Skp. To design such a variant, I first inspected crystal structure 1SG2 (figure 5.3) for residues in Skp that make contacts at the trimerization interface. Next, I designed one construct, FAI (for Full Alanine Interface) with all of these residues (figures 5.3 and 5.4) mutated to alanine. In a second approach, I elected to target one residue in the same interface that was conserved. Sequence alignment (figure 5.4), revealed that Y124 (Y136 in crystal structure with PDB ID 1SG2, figure 5.3) shows conservation. To target this residue I designed a variant Y124A that I named SAI (for Single Alanine Interface). Both FAI and SAI expressed and purified similarly to wild-type Skp (figures 5.5 and 5.6A-D). Notably, FAI eluted at a later time than Skp in size-exclusion chromatography, consistent with weakened trimerization (figure 5.5).

Next I set out to verify using CD that the FAI and SAI constructs were correctly folded. Consistent with native-like structure, both constructs show α-helical CD

signatures like Skp (figure 5.6). Furthermore, sedimentation equilibrium (SE) was achieved within a similar time scale as Skp and remained stable throughout the SE runs.

SE of FAI or SAI showed that trimerization free energies varied as might be expected for the two constructs, in the pattern $\Delta G°_{WT} < \Delta G°_{SAI} < \Delta G°_{FAI}$ (figure 5.7A). Importantly, FAI SE data fit to a single species, showing that FAI does not form trimers. Therefore, FAI is a mostly folded monomeric form of Skp. This provides further evidence that monomeric Skp does not completely unfold (Burmann, Holdbrook et al., 2015, Sandlin, Zaccai et al., 2015). For SAI (Y124A), binding energy was reduced by 1.8 kcal mol^{-1} (n=1), or about 0.6 kcal mol^{-1} Tyr^{-1}. Thus, conservation of this residue is partly due to its importance for trimerization (figure 5.7A). This importance may be due to its position in Skp, where it may act like a "strut" locking each long helix its neighbor (figure 5.7B).

Skp Trimerization is Necessary for its Antiaggregation Activity (AA)

With FAI in hand, I addressed the switching hypothesis by comparing the AA of Skp with that of FAI. To make this comparison, I used SV to quantify AA (see section 5.2). As shown in figure 5.9, at 40 minutes post-dilution (mpd), Skp has a robust AA with uOmpLa in 150 mM NaCl, 300 mM urea and pH 8.0. On the other hand, FAI has a much lower AA than Skp (figure 5.9). This finding supports the switching hypothesis and the idea that Skp trimers "cage" uOMPs to provide AA (Walton, Sandoval et al., 2009).

Skp Trimers May be Necessary for OMP Binding

One explanation for this loss of AA is that binding to FAI is weakened. To test this idea, I performed thermodynamic measurements using our intrinsic tryptophan

fluorescence assay (Moon, Zaccai et al., 2013) with the FAI construct. Preliminary data in figure 5.8 suggests that loss of trimerization could weaken binding dramatically, with $\Delta G°_{FAI}$ = -5.4 ± 0.1 kcal mol^{-1} at 25 °C. This result is consistent with the cage formation observed in NMR and SANS solution structures (Burmann, Wang et al., 2013).

Skp Antiaggregation Activity (AA) Differs with Each uOMP Client and Does not Depend on Size

Skp is believed to assist SurA and FkpA within the periplasmic chaperone network (Sklar, Wu et al., 2007). To provide this assistance, Skp must have AA with many different uOMPs. However, as a cage for uOMPs, Skp has a limited size. Therefore, I set out to test whether Skp AA decreases with uOMP size. To determine how AA depends on size, I used my SV AA assay with uOmpA, uOmpLa, and uBamA-TM, which have molecular weights (M_W) of 35.30, 30.97, and 43.67 kDa, respectively. As shown in figure 5.9A, Skp AA varies between the three uOMPs. Still, neither chain length nor M_W explain AA (figures 5.10G and 5.10K).

Surprised by this, I estimated the R_H of uOMPs from their sedimentation coefficients ($S_{20,W}$) (Teller, 1973). The $S_{20,W}$ values for uOmpLa and uBamA-TM were estimated using linear extrapolation of $S_{20,W}$* as a function of urea (section 5.2, table 5.2, and figure 5.11A). For uOmpA, I used Danoff's published value of 2.35 S (Danoff and Fleming, 2011). To calculate the R_H of uOMPs, I employed the program SEDNTERP's Teller method (Teller, 1973). As expected (Flory, 1953), the R_H of uOMPs does not explain Skp AA any better than M_W (figure 5.10H).

This lack of correlation led me to examine the Skp AA for uOmpA more closely, because is the lowest among the three OMPs (figure 5.9A). This AA is lowest because the numerator (section 5.2) is small. The small numerator is related to uOmpA's slow aggregation rate (AR) (Danoff and Fleming, 2011). This simple observation shows that uOmpA, with slow AR, cannot benefit from Skp AA.

Antiaggregation Activity (AA) is Not Correlated with Parameters Calculated from uOMP Primary Structure, Including a Novel Hydrophobic Segregation Parameter

With the knowledge that AR partly dictates AA, I sought to further explain AA by examining the influence of amino acid composition. No parameter based on amino acid composition correlated with AA (figure 5.10). Initially, the sequence density of tryptophan showed correlation (figure 5.10E). However, a negative control, uOmpLa with all its tryptophan residues mutated to phenylalanine (OmpLa-WF), showed a similar AR as the wild-type uOmpLa (figure 5.9C, dotted line, is representative).

Kappa (κ) is a score describing the segregation of charge along a polypeptide chain (Das and Pappu, 2013). However, κ did not correlate with AA (figure 5.10H). Having exhausted other parameters, I devised a hydrophobic asymmetry parameter, analogous to κ (see section 5.2). Naming the parameter ξ, I tested whether ξ explains AA by calculating ξ for all three uOMPs. Figure 5.10J shows that hydrophobic segregation does not correlate with AA either.

Skp is Not a Holding Chaperone for uOMPs

For Skp to be a holding chaperone, its AA must correlate with its CA (section 1.2 and figure 1.5). Entzminger et al. showed Skp has holding CA for antibody fragments (Entzminger, Chang et al., 2012). In contrast, Thoma et al. (2015) found that Skp did not have CA for FhuA folding. Thus, I set out to determine whether Skp is, in fact, a holding chaperone for uOMPs. Using conditions identical to those used to measure AA (20 mM Tris-HCl, 150 mM NaCl, 300 mM urea, pH 8.0), I conducted delayed folding assays (figure 3.1) using diC11PC LUVs.

Skp shows no CA for OmpA, OmpLa, or BamA-TM (figure 5.14). For OmpA, Skp actually inhibited folding in these conditions, as seen before by others (Bulieris, Behrens et al., 2003, Patel and Kleinschmidt, 2013). As a control for CA, I tested full-length BamA, for which Skp did show CA. Therefore in this case I find that Skp is not a holding chaperone for uOMPs, despite its ability to bind uOMPs (Qu, Mayer et al., 2007, Moon, Zaccai et al., 2013) and display AA with uOmpLa and uBamA-TM (figure 5.9).

uOmpLa and uBamA Aggregation Occurs at Nanomolar Concentrations

In the periplasm, uOMPs may bind SurA (Wu, Ge et al., 2011, Thoma, Burmann et al., 2015), preventing free uOMPs from reaching the 2 μM concentrations used in these assays. If so, uOMP aggregation might not be biologically relevant. Therefore, I sought to estimate the lowest concentration at which uOMPs can aggregate. To identify such a low concentration threshold I developed a new sensitive assay using intrinsic tryptophan fluorescence.

To develop this assay I first verified that fluorescence kinetics depend on protein and urea concentration. Figure 5.12A shows that rapid dilution of uOmpLa stock produces a tryptophan fluorescence signal (Ex. 295 nm, Em. 341 nm) that increases with time. As expected for aggregation, these kinetics depend on protein (figure 5.12A) and urea (figure 5.12C) concentration.

To verify the assay, I also measured aggregation kinetics by SV. Figure 5.12D shows the concentration dependence of uOmpLa aggregation kinetics in 4 M urea. At 198 nM (figure 5.12D, lowest concentration), SV shows that uOmpLa is monomeric in 4 M urea. At 198 nM uOmpLa and 4 M urea, fluorescence kinetics do not increase with time (figure 5.12C). Thus, both methods show similar dependence on urea and uOmpLa concentration. This shows that tryptophan fluorescence monitors uOmpLa aggregation. However, fluorescence and aggregation have a different mechanism (figure 5.12E). Fluorescence increases more slowly than aggregation in the first few minutes, and faster than aggregation after a few hours (figure 5.12E). However, both processes reach steady state after 24 hours.

Because fluorescence is proportional to aggregate at equilibrium, a plot of fluorescence amplitudes versus uOmpLa concentration (figure 5.12B) is a straight line. Extrapolating this line to zero reveals that, in 150 mM NaCl, 80 mM urea and pH 8.0, uOmpLa aggregates above 18.9 ± 25.5 nM. The same method used with BamA and 2 mM TCEP (figure 5.13) estimates an aggregation threshold of 44.0 ± 37.4 nM. These numbers are very close to the affinity of Skp, which is in the nanomolar range (Moon, Zaccai et al., 2013).

5.4 Discussion

In chapter 4, I suggested a switching hypothesis that predicted that Skp trimers are the active form of Skp. Consistent with the switching idea, a variant of Skp that does not form trimers (figure 5.7) has reduced AA (figure 5.9). This reduced AA correlated with reduced affinity to the FAI monomer (figure 5.8). These findings support the switching model: Skp trimers are upregulated by uOMPs to provide high binding affinity and AA, while monomeric Skp is inactive (Sandlin, Zaccai et al., 2015).

Skp AA also varies between uOMPs (figure 5.9), as it does with antibody fragments (Entzminger, Chang et al., 2012). However, out of three uOMPs studied, uOmpA behaved uniquely (figure 5.9). The weak Skp AA for uOmpA is a consequence of the slow AR of uOmpA (Danoff and Fleming, 2011). This suggests that OmpA is part of a class of "soluble" OMPs like OmpX (Ebie Tan, Burgess et al., 2010) that bind Skp (Burmann, Wang et al., 2013), but do not benefit from its AA. Oddly, these "soluble" OMPs are highly expressed (Masuda, Saito et al., 2009), alongside "insoluble" OMPs like OmpLa and BamA (figures 5.11 and 5.12). This raises an interesting question about whether "soluble" uOMPs compete with "insoluble" uOMPs for Skp AA in *E. coli*. Unfortunately, AA could not be explained by uOMP primary structures (figure 5.10). However, my data are limited by extended dead-times (20 seconds for fluorescence, 20-30 minutes for SV). In the future, stopped-flow kinetics might better explain the link between AA and protein structure.

Whereas Enztminger et al. (2012) concluded that Skp is a holding chaperone, I found that this is not the case for native uOMP substrates. For uOMPs folding into membranes, neither binding nor Skp AA are good predictors of CA. For OmpA, binding

occurred, AA was insignificant, and folding was inhibited. Such inhibition of OmpA folding by Skp was observed before (Bulieris, Behrens et al., 2003, Patel and Kleinschmidt, 2013). Bulieris et al. and Patel et al. showed, respectively that addition of LPS and BamA enables Skp CA for uOmpA. However, LPS was shown to denature Skp (Burmann, 2015), and Skp could have somehow stimulated the BamA foldase activity (Gessmann, Chung et al., 2014, Plummer and Fleming, 2015). On the other hand, for uOmpLa and uBamA-TM, Skp provided AA, but had no effect on folding.

A new sensitive fluorescence assay measures a component of uOMP aggregation that involves the tryptophan side chains (figure 5.12E). It may be that uOmpLa aggregates are amyloids, like uOmpA (Danoff and Fleming, 2015), with structures similar to α-synuclein that buries rows of aromatic side-chains (figure 1.4) (Tuttle, Comellas et al., 2016). Regardless, this new assay revealed that uOmpLa and uBamA will aggregate if even 1% of their periplasmic concentrations (~μM) are free (Masuda, Saito et al., 2009). This is 100 times lower than the threshold for uOmpA aggregation (Ebie Tan, Burgess et al., 2010, Danoff and Fleming, 2011). Such a low threshold for aggregation, without disaggregase activity to recover from it (see chapter 1), means free uOMP concentrations must be tightly regulated. Such regulation is uniquely provided by Skp, which possesses nanomolar binding affinity for uOMPs (Moon, Zaccai et al., 2013).

To summarize chapter 5, Skp is an excellent AA factor for uOMPs, whereas its holding chaperone activity is weak and poorly understood. The Skp trimer is the active form of the protein that is switched on by unfolding stresses. Thus, the major function of Skp in periplasmic proteostasis is not as a chaperone for OMPs, but as a “protein buffer”:

Skp kinetically and thermodynamically absorbs free uOMPs to keep them below dangerous nanomolar concentrations.

5.5 Tables and Figures

Table 5.1. Calculated ξ Values for uOMPs in This Study and Commonly Used in Our Experiments.

Name	ξ (Unitless)	Name	ξ (Unitless)
OmpA-TM	0.2560	BamA-TM	0.2459
OmpA	0.2268	BamA	0.2130
OmpX	0.3023	OmpLa	0.2344
OmpW	0.2342	OmpF	0.2175
FadL	0.1941	OmpT	0.2377
PagP	0.3439	OmpG	0.2145

Table 5.2. Hydrodynamic Parameters and AA for Proteins.

Name	SV Temp °C	Time Rotor Start[1] (min)	NaCl (M)	Urea (M)	N[2]	C_{load}[3]	C_{mon}[4]	ΔC_{mon}No Skp[5]	ΔC_{mon} With Skp[6]	AA[7]	$S_{20,W}$*[8] (Svedbergs)	M_W[9] (kDa)	WSSR/ DoF[10]	R_H[11] (Å)
OmpLa Extrapolated	20	30	0	0.0*	10	-	-	-	-	-	2.26 ± 0.17	31.3 ± 0.071	8.10 ± 0.50E-05	31.5 ±1.2
BamA-TM Extrapolated	20	30	0	0.0*	6	-	-	-	-	-	2.59 ± 0.051	25.5 ± 0.93	2.46 ± 0.70E-05	38.7 ±0.8
OmpA	25	30	0.15	0.3	3	0.092 ± 0.004	0.074 ± 0.002	-0.017 ± 0.003	-	-	2.73 ± 0.019	*35.30*	4.37 ± 1.41E-06	35.8
OmpA / Skp	25	30	0.15	0.3	3	0.142 ± 0.005	0.122 ± 0.004		-0.019 ± 0.004	0.89 ± 0.43	5.28 ± 0.018	*85.19* (66.6±3.54)	1.62 ± 1.85E-04	46.1 ±0.0
OmpLa	20	30	0.15	0.3	3	0.207 ± 0.001	0.020 ± 0.02	-0.187 ± 0.001			*2.25*	*30.97*	6.22E-06 ± 0.15E-06	-
OmpLa / Skp	25	30	0.15	0.3	3	0.213 ± 0.006	0.161 ± 0.005		-0.052 ± 0.009	3.56 ± 0.87	5.72 ±0.037	*80.86* (64.0±2.29)	7.91 ± 0.35E-06	32.8 ±0.2
BamA-TM	25	30	0.15	0.3	3	0.251 ± 0.008	0.067 ± 0.002	-0.181 ± 0.010			*2.56*	*43.67*	1.36 ± 0.10E-04	-
BamA-TM / Skp	25	30	0.15	0.3	3	0.180 ± 0.012	0.087 ± 0.008	-	-0.098 ± 0.007	1.85 ± 0.31	4.99 ±0.024	*93.55*	9.58 ± 0.45E-05	49.5 ±0.2
OmpLa	25	40	0.15	0.3	2	0.216 ± 0.011	0.020 ± 0.002	-0.196 ± 0.001			*2.25*	*30.97*	5.80 ± 0.72E-06	-
OmpLa / Skp	20	40	0.15	0.3	2	0.212 ± 0.010	0.137 ± 0.003		-0.076 ± 0.010	2.59 ± 0.31	4.91 ±0.017	*80.86* (57.36±1.56)	1.38 ±0.15E-05	-
OmpLA / FAI	25	40	0.15	0.3	2	0.218 ± 0.029	0.030 ± 0.004		-0.188 ± 0.034	1.04 ± 0.24	N/A	*47.60*	8.25 ±0.14-06	-

1. The time after dilution of uOMP from high urea. Also, the time of rotor start.
2. The number of experiments used to provide the estimates.
3. The three OD280 measurements taken from the first scan of each treatment and closest to 6.5 cm were averaged over all replicates. Error is the SEM.
4. C0, the absorbance contribution of the late-sedimenting species was estimated by fitting to a single Gaussian (see section 5.2). The figure quoted is the mean of the measurements. Error is propagated as the square root of the sums of squares of the errors for the individual fits.
5. The estimate and propagated error for C_{mon}-$C_{load,}$ if no Skp was present.

6. The estimate and propagated error for C_{mon}-$C_{load,}$ if Skp was present.
7. Antiaggregation activity, as defined in the methods for SV. This is calculated as $<C_{mon}\text{-}C_{load}>$ without Skp / $<C_{mon}\text{-}C_{load}>$ with Skp. Error for a ratio of means is propagated using the Taylor expansion method, according to the formula $\sqrt{\mathrm{Var}\left(\frac{\bar{X}}{\bar{Y}}\right)} \approx \sqrt{\frac{1}{\bar{Y}^2}s_X^2 + \frac{\bar{X}^2}{\bar{Y}^4}s_Y^2 - 2\frac{\bar{X}}{\bar{Y}^3}s_{XY}}$, where s_x and s_y are the respective sample standard deviations of X and Y, and S_{XY} is the sample covariance of X and Y.
8. The observed sedimentation coefficient from a single Gaussian fit using DCDT+, or extrapolated to zero denaturant in the case of uOMPs. Due to poor fits to single Gaussians for uOmpLa and uBamA-TM in the AA tests without Skp, in these treatements $S_{20,W}^*$ was fixed in the regression to the value estimated by extrapolating to 300 mM urea *(underlined italics).*
9. Either the fixed, calculated M_W *(underlined italics)*, or the fitted value estimated from a fit of the late-sedimenting species to a Gaussian profile using DCDT+. If a second number is included in parentheses (e.g., "(64.0±2.29)"), it is the estimate if the M_W is allowed to float in the regression.
10. Weighted Sums of Squares of Residuals divided by the number of data points for the Gaussian fit to the late-sedimenting species, as reported by the DCDT+ software. This value is averaged over all replicates with the indicated SEM. The theoretical WSSR was chosen because it provides a better estimate of the overall error between experiments. This is because g(S*) distributions have error that decreases with increasing S*.
11. For uOmpLa and uBamA-TM, the apparent hydrodynamic radius (R_H), calculated using the observed $S_{20,W}^*$ listed in the same row. For OmpA, the published value for $S_{20,W}^*$ to estimate R_H (Danoff and Fleming, 2011). The error was omitted in this case.

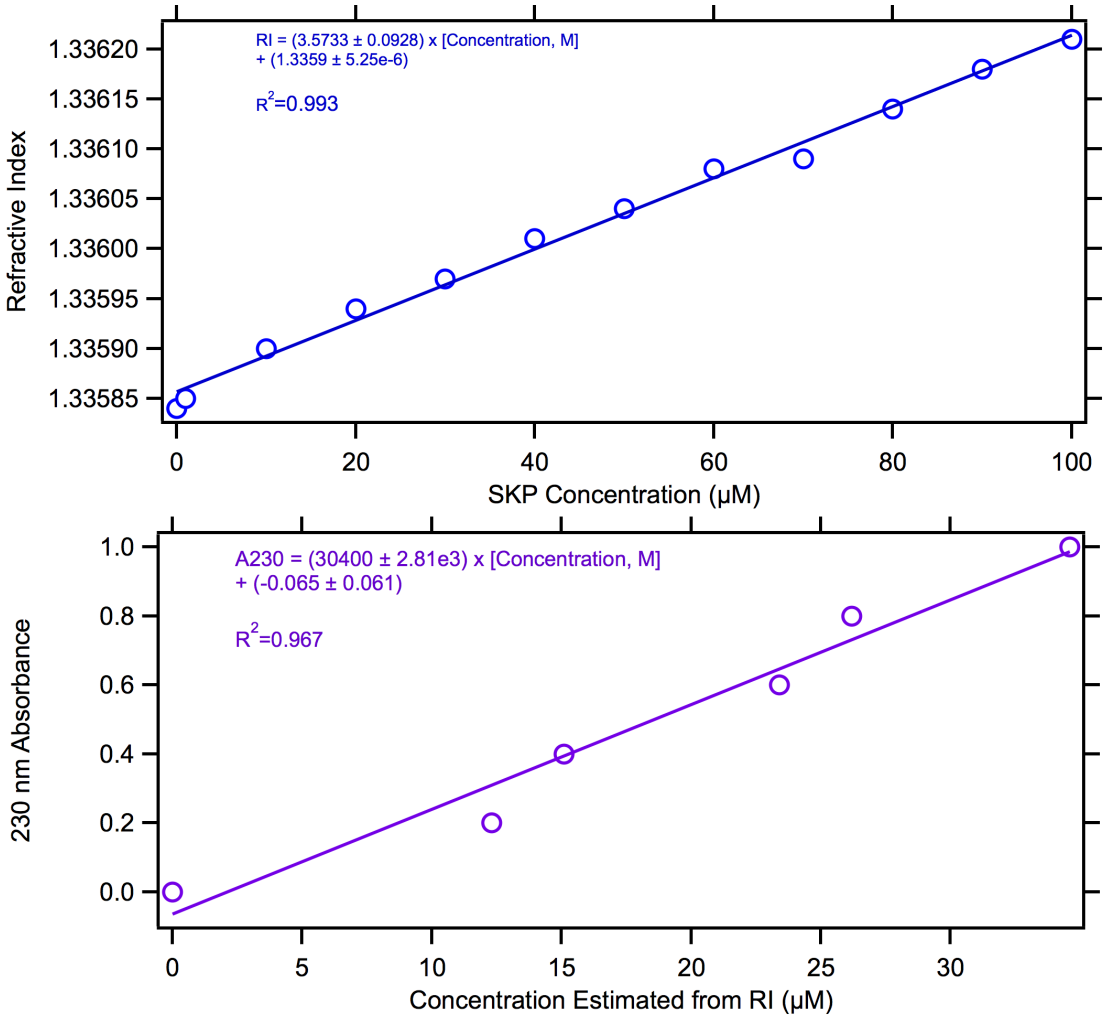


Figure 5.1. Differential Refractometry Used to Estimate Skp Variant Concentrations. Because the Bradford assay gave spurious results with Skp, the extinction coefficient of SAI was estimated using differential refractometry. A., Plot of refractive index versus Skp concentration as estimated from absorbance. B., Plot of absorbance at 230 nm for SAI versus SAI concentration estimated from the RI of SAI and the regression in A. The slope of the regression in B. gives ϵ_{230}=30400 M^{-1} cm^{-1}. All measurements were performed in 20 mM Tris-HCl, 200 mM NaCl pH 8.0 at 20 °C. The extinction coefficient of SAI was also used for FAI.

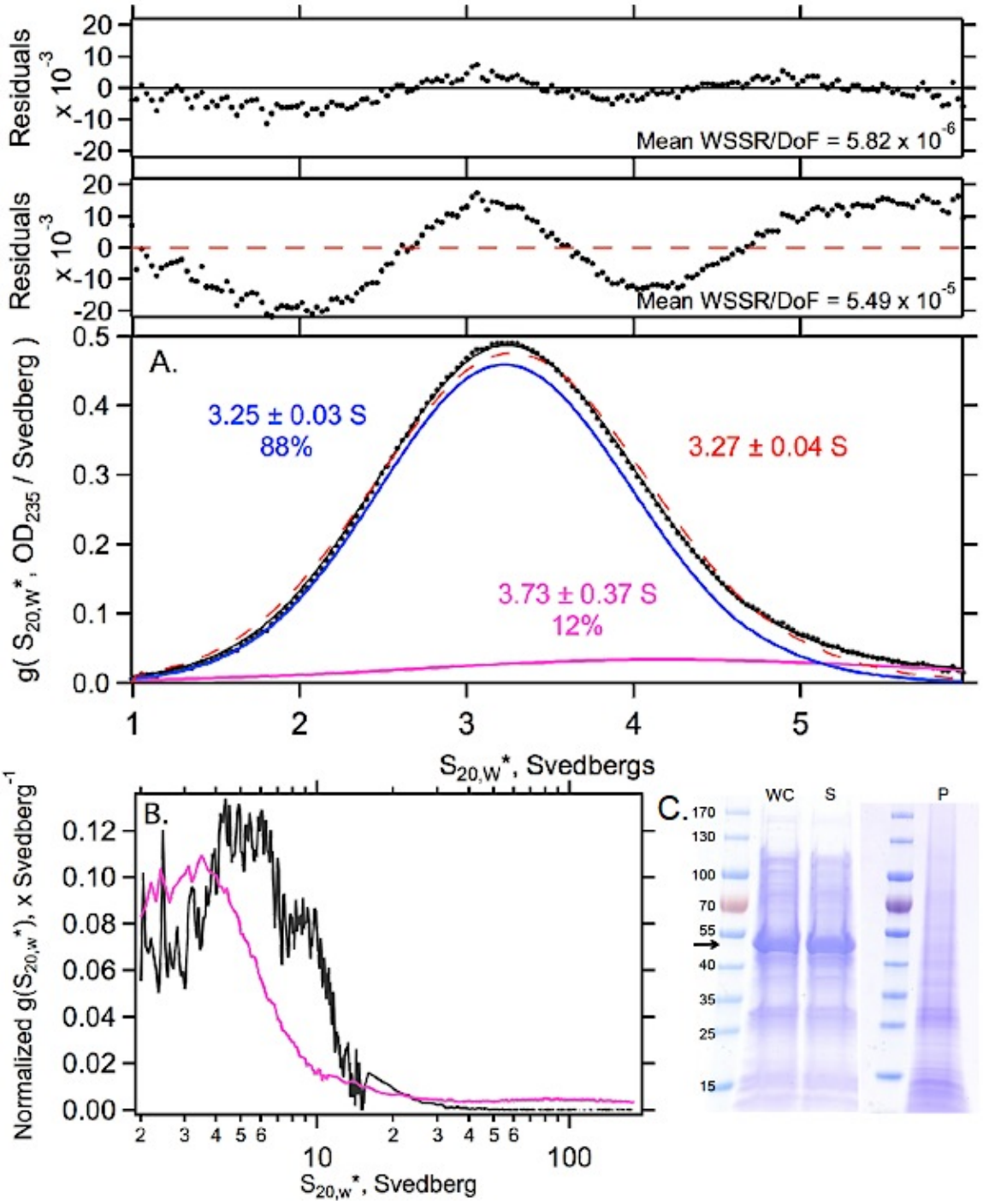


Figure 5.2. Skp Sediments as Two Species, and BamA-TM Forms Low Molecular Weight Aggregates in Inclusion Body Preparations. A., Skp g($S_{20,w}$*) curve generated from SV data collected at 25 °C in GF buffer (20 mM Tris-HCl, 200 mM NaCl, pH 8.0). Black line, fit to two species with molecular weights fixed to that of the Skp monomer

(16.65 kDa) and trimer (49.94 kDa). Dotted red line, fit to a single species with all parameters allowed to float. Blue, trimer, magenta, monomer. Figures indicate the estimated sedimentation coefficients and percent contributions to absorbance. B., normalized $g(S_{20,W}*)$ curves generated from SV data. Black, BamA-TM in GF buffer forms aggregates with S* ~2-13. Magenta, BamA-TM after chloroform-methanol extraction (see section 5.2), at 30 minutes post-dilution from 6 M urea into 20 mM Tris-HCl, 150 mM NaCl, 300 mM urea, 2 mM TCEP, pH 8.0. C., Purification gels showing BamA-TM partitions to the soluble fraction in lysis buffer. Arrow, BamA-TM. WC, whole cells, S, soluble fraction, P, insoluble fraction.

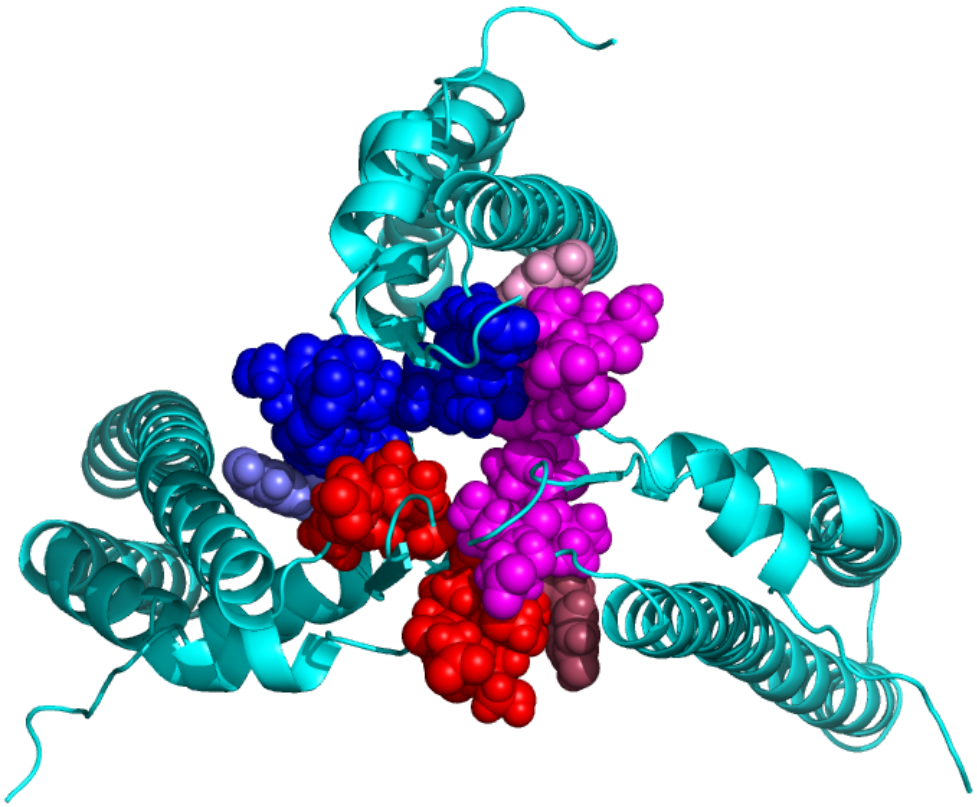


Figure 5.3. Residues Chosen For Alanine Mutation Due to Their Positions in the Skp Trimerization Interface. The residues chosen for alanine mutation within each protomer in the Skp trimer are indicated in different colors (magenta, blue, and red). Residue Y124 is shown as a lighter color (pink, slate, and raspberry). Model from Zaccai, Sandlin et al. (2016).

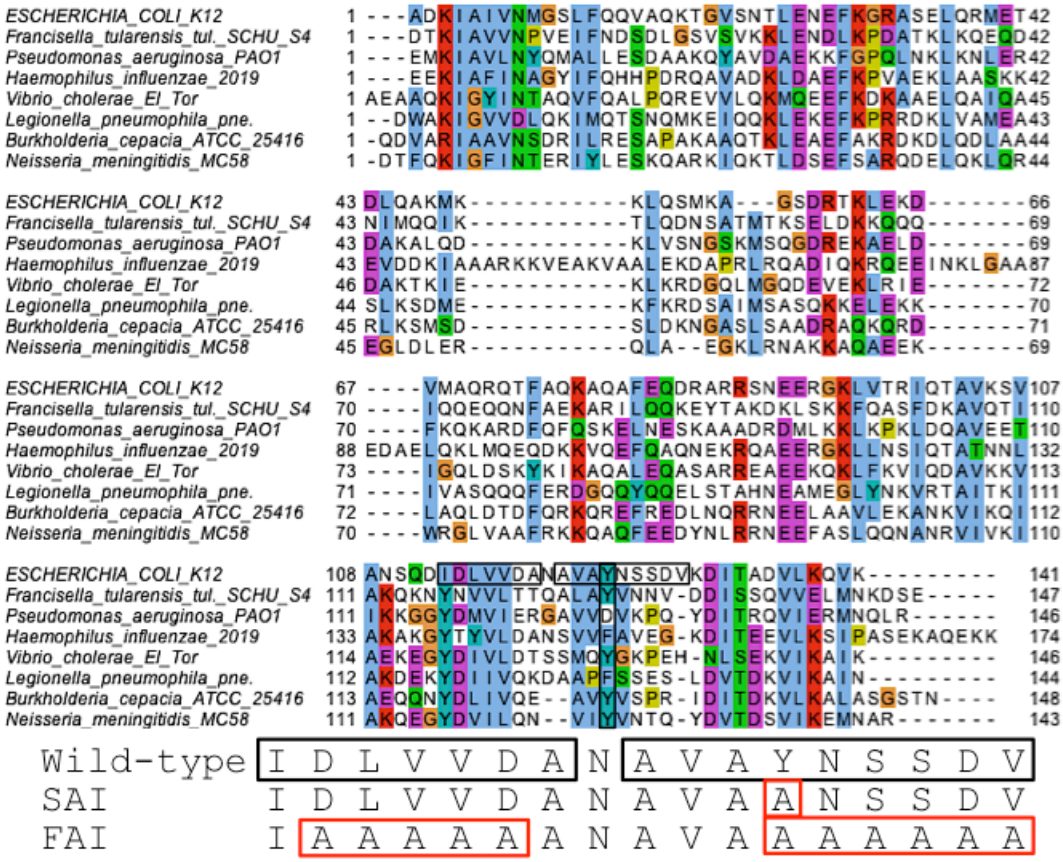


Figure 5.4. Sequence Conservation of Skp and Skp Alanine Variants. Top, Clustal Omega (Sievers, Wilm et al., 2011) global alignment of representative Skp sequences detected by PSI-BLAST (Altschul, Madden et al., 1997) in the β- and γ-proteobacteria. No examples of Skp-like genes were detected by PSI-BLAST within the α- or ε-proteobacteria, or the orders *chlamydiales*, *spirochaetales*, *bacteroidales*, or *fusobacteriales*. This shows that Skp is not found in all Gram-negative bacteria. Black boxes show the residues shown as spheres in figure 5.3 and the position of Y124 in the alignment. Bottom, alignment of Skp with the alanine variants SAI and FAI. Black boxes, residues shown as spheres in figure 5.3. Red boxes, alanine mutations in SAI and FAI variants. The alignment image was created with the program Jalview.

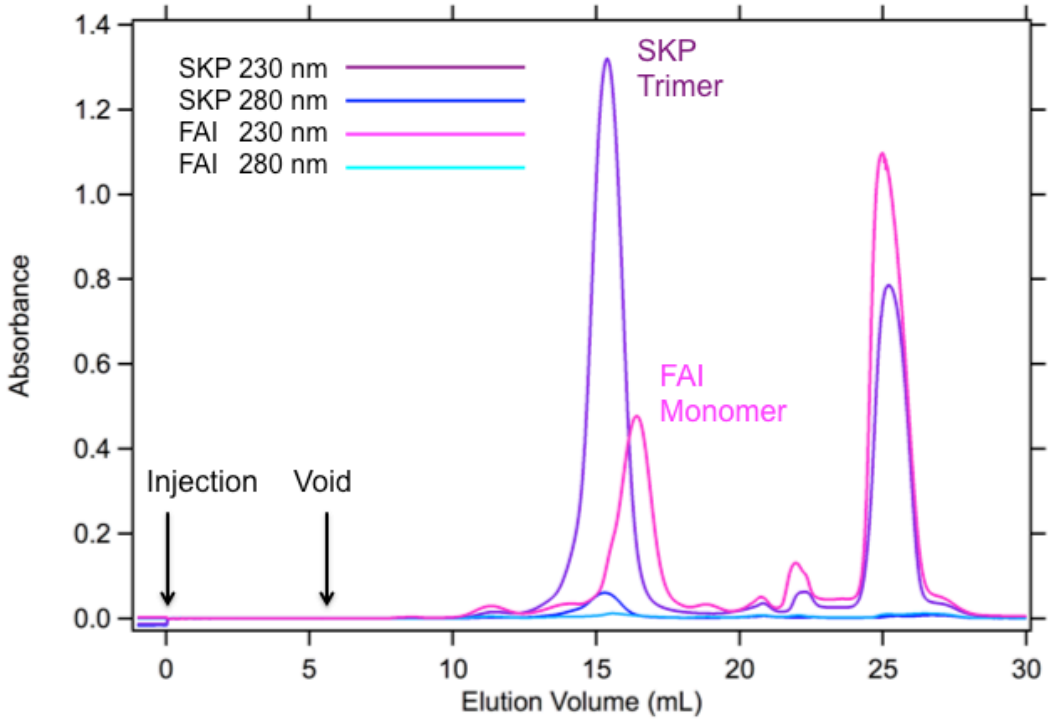


Figure 5.5. Typical Size-Exclusion Chromatography of Skp and Skp-FAI Shows FAI is Monomeric. Purple, the major peak near 15 mL represents the Skp trimer. Magenta, the major peak for FAI, elutes later than Skp. Note that the FAI OD280:OD230 (cyan:magenta) is reduced relative to the Skp OD280:OD230 ratio (blue:purple), consistent with the Y124A mutation (see figures 5.3 and 5.4). Injections were 1 mL samples in elution buffer containing 50 mM Tris-HCl, 0.5 M NaCl, 262.5 mM imidazole. The exchange buffer is 20 mM Tris-HCl, 200 mM NaCl (pH 8.0).

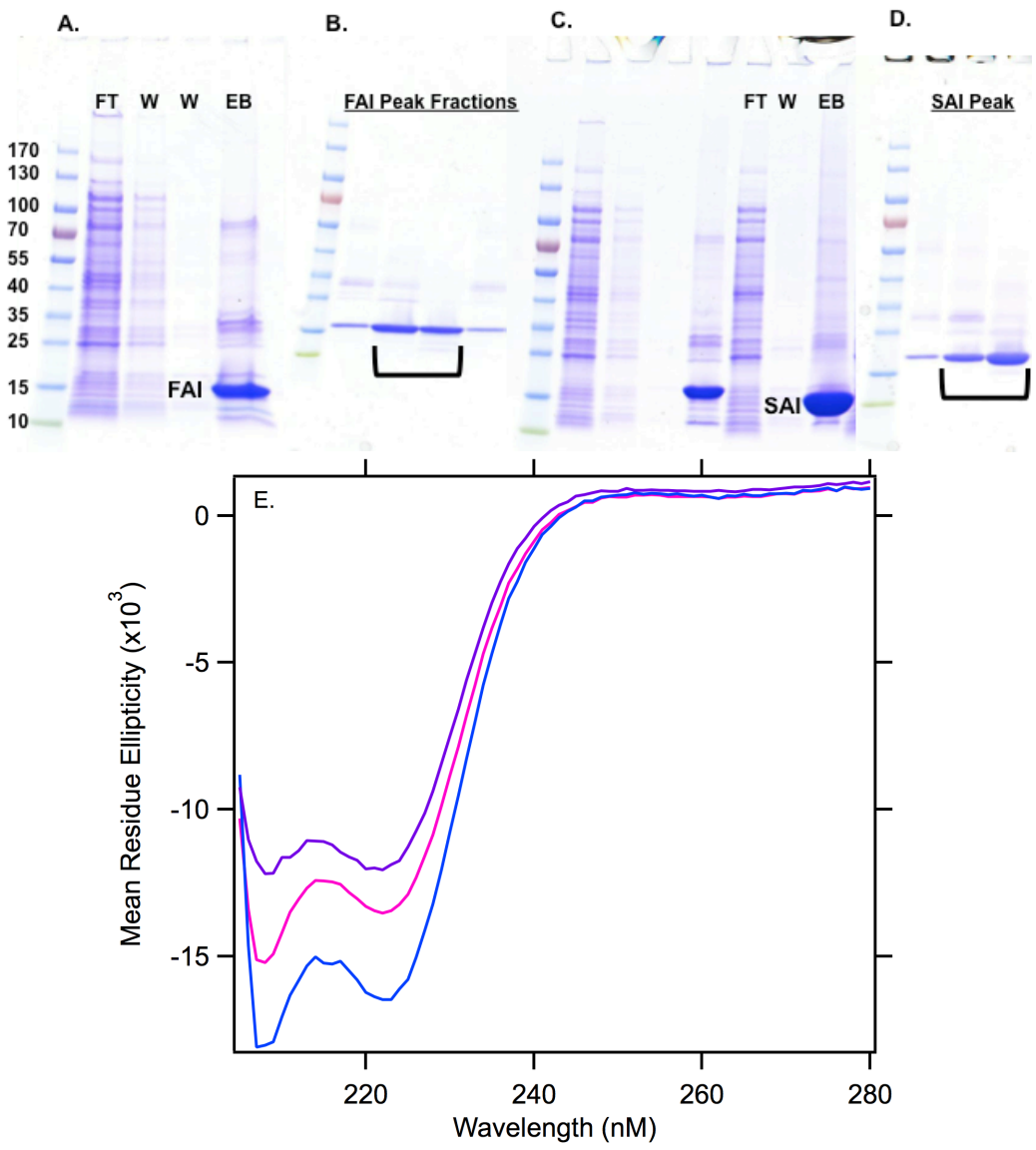


Figure 5.6. FAI and SAI are Pure and Show α-helical CD Signatures. A., Nickel-affinity purification (NAP) of FAI. Lanes 2-5, flow-through, wash (5 mL), wash (3 mL), and elution. B., Fractions from the major peak from size-exclusion chromatography (SEC) of FAI (see figure 5.5). Brackets, fractions taken. C., NAP. Lanes 2-5, the unused variant. Lanes 6-8, flow-through, wash, and elution of SAI. D., Fractions from the major peak from SEC of SAI. E., CD spectra for Skp (blue), FAI (magenta) and SAI (purple), show the α-helical signature of a folded Skp structure.

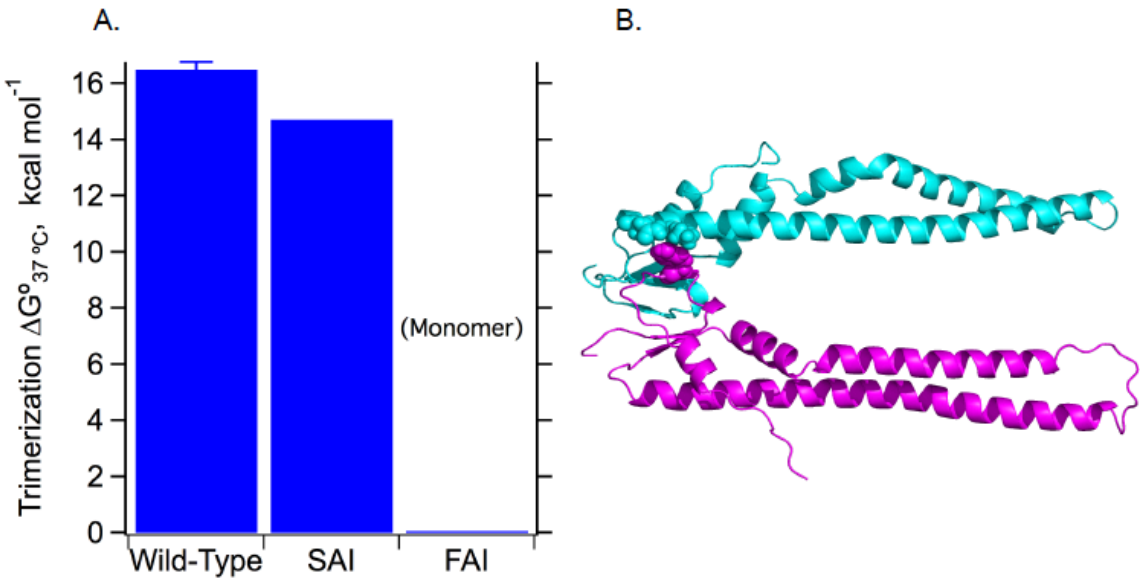


Figure 5.7. Residues Mutated to Alanine in FAI and SAI Variants Stabilize the Skp Trimer. A., the trimerization energetics of three Skp variants estimated by SE at 26 °C in 20 mM Tris-HCl, 200 mM NaCl, pH 8.0. Y124A results in a 1.8 kcal mol^{-1} reduction in trimerization energy, while the FAI variant (see figures 5.3 and 5.4) results in a monomer. The square-roots of variances (SRV) for the fits in A., normalized by their degrees of freedom (DoF), are as follows: Skp (from table 4.3), 4.80×10^{-6}, SAI, 4.20×10^{-6}, FAI, 4.37×10^{-6}. B., Magenta and cyan, two adjacent protomers in Skp. Y124 (magenta spheres) stabilizes the Skp trimer by interacting with K106 and N110 (cyan spheres).

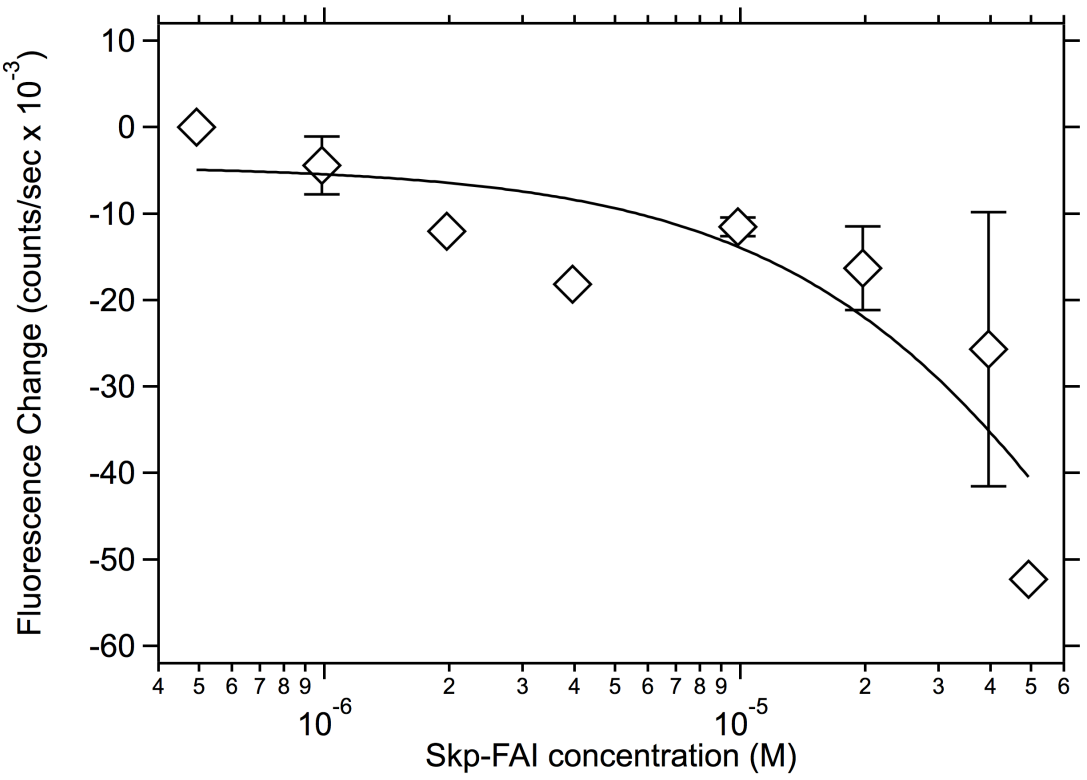


Figure 5.8. Preliminary Data Suggest FAI Binds uOmpLa in the Micromolar Range. Fluorescence binding data for FAI to uOmpLa collected at 25 °C. Final concentrations were 40 nM uOmpLa, 20 mM Tris-HCl, 50 mM NaCl, 80 mM urea and pH 8.0. The black trace is a fit to a single-site binding model (Moon, Zaccai et al., 2013), The estimated K_D from the fit is 120 ± 18.3 μM. Diamonds with error bars, n=2, diamonds without error bars are singlet data.

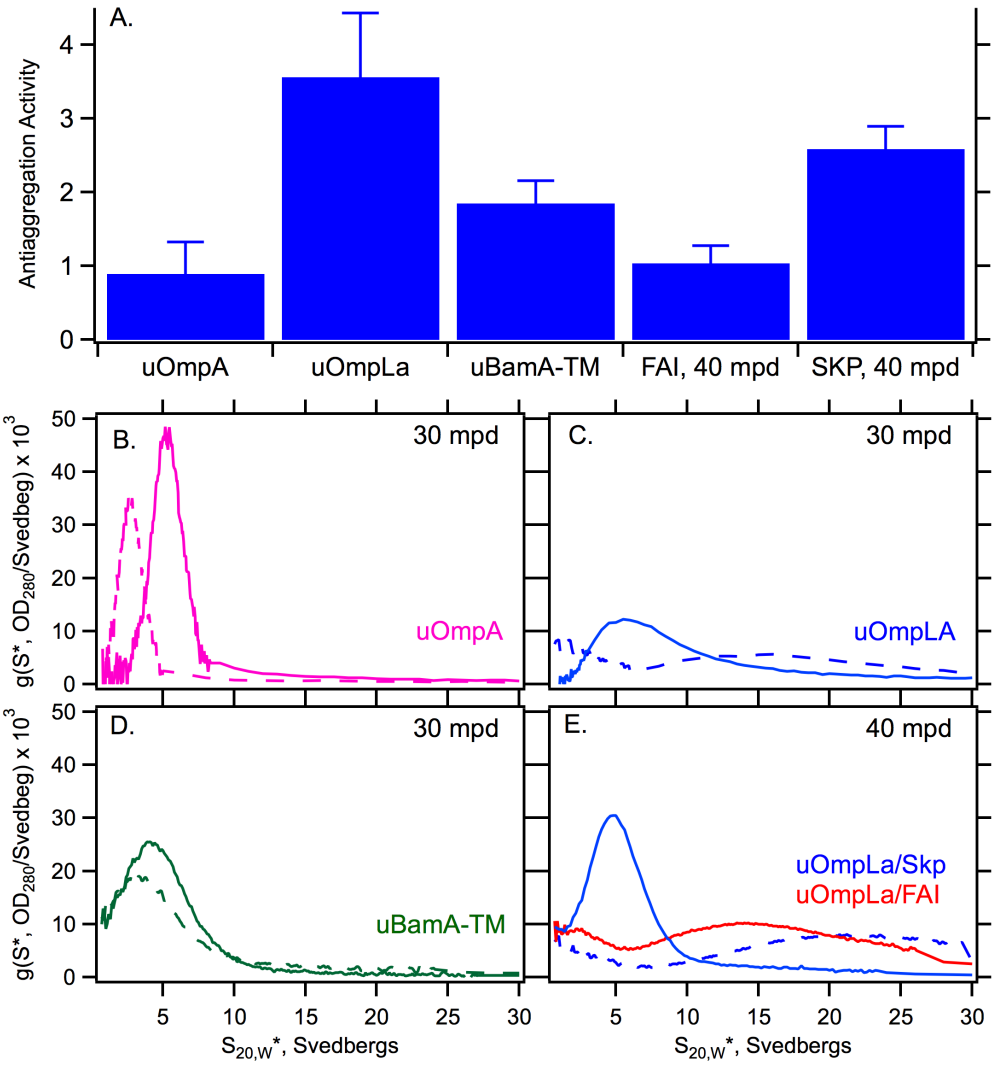


Figure 5.9. Skp Trimers Provide Antiaggregation Activity (AA) for uOmpLa and uBamA-TM. A., Left three bars, Skp AA varies with three uOMPs at 30 minutes post-dilution (mpd). Right two bars, AA of 26.8 μM FAI is insignificant compared to 18 μM Skp at 40 mpd. B-E, SV g($S_{20,W}$*) plots related to A, collected at 280 nm. Solid lines, with Skp, dotted lines, without Skp. B., uOmpA. C., uOmpLa*. D., uBamA-TM. E., blue, uOmpLa with and without 18 μM Skp. Red, SV data for uOmpLa with 26.8 μM FAI*. Asterisks (*) mean that SV temperatures varied from 20-25 °C between conditions (see table 5.2).

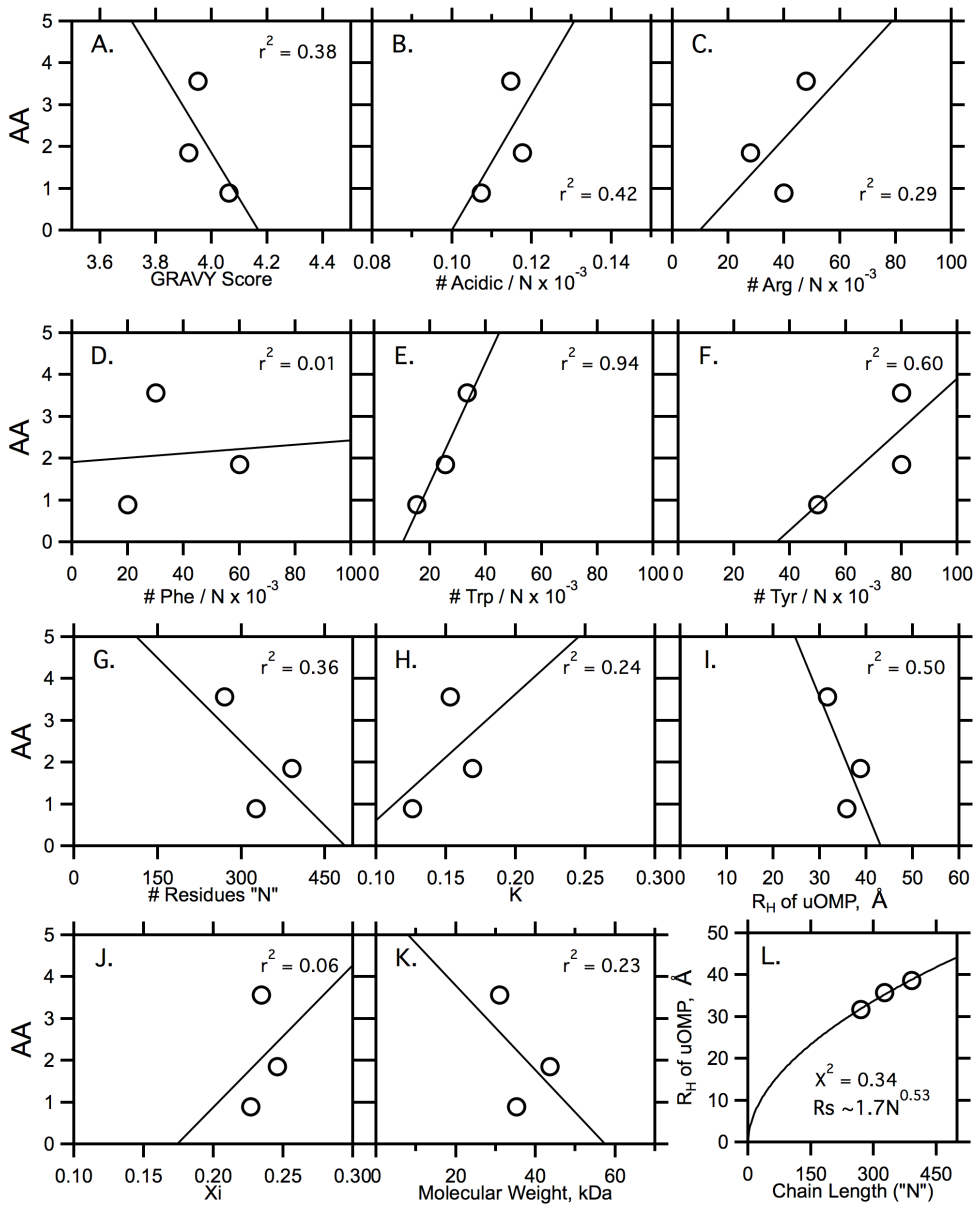


Figure 5.10. Antiaggregation Activity is Not Explained by Parameters Calculated from the Amino Acid Sequence of uOMPs. A.-D. Hydropathy, sequence density of acidic residues, arginine, and phenylalanine do not explain AA. Tryptophan density, E., seemed to explain the AA. However, uOmpLa-WF with all tryptophans mutated to phenylalanine aggregated at the same rate as wild-type uOmpLa (figure 5.9C, dotted blue is representative). F.-K., sequence density of tyrosine, chain length, Kappa (κ) (Das and

Pappu, 2013), apparent R_H estimated by SV (Danoff and Fleming, 2011) (also table 5.3), Xi (ξ) (see section 5.2 and table 5.1), and molecular weight all explain AA poorly. L., Estimated R_H values for uOMPs are physically reasonable. A fit to a Flory expression of the form $R_H = R_0N^v$ (Flory, 1953, Wilkins, Grimshaw et al., 1999), with R_0=1.66 and v=0.53 is underdetermined (χ^2=0.34) but consistent with estimates of R_H from SV.

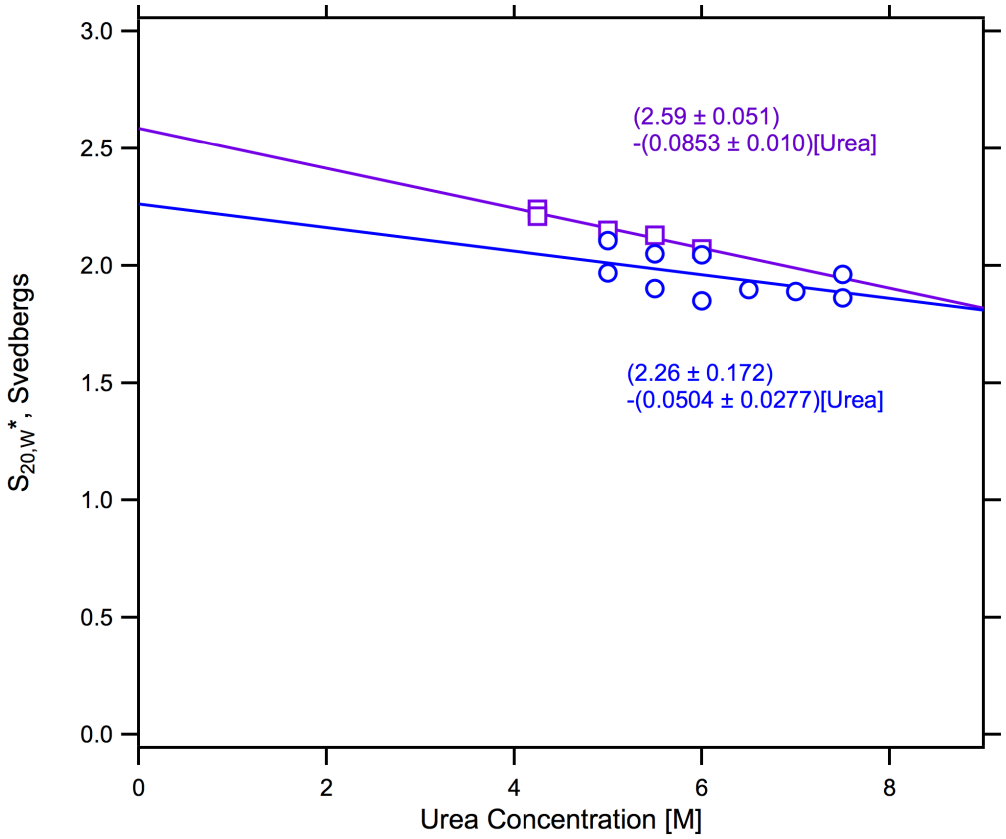


Figure 5.11. Urea Extrapolation Used to Estimate $S_{20,W}$* for uBamA-TM and uOmpLa. Observed $S_{20,W}$* as a function of urea concentration for uBamA-TM (purple) and uOmpLa (blue). Extrapolation to zero denaturant by linear regression is used to estimate $S_{20,W}$* values in the absence of denaturant and for fixing parameters to estimate AA (see table 5.2).

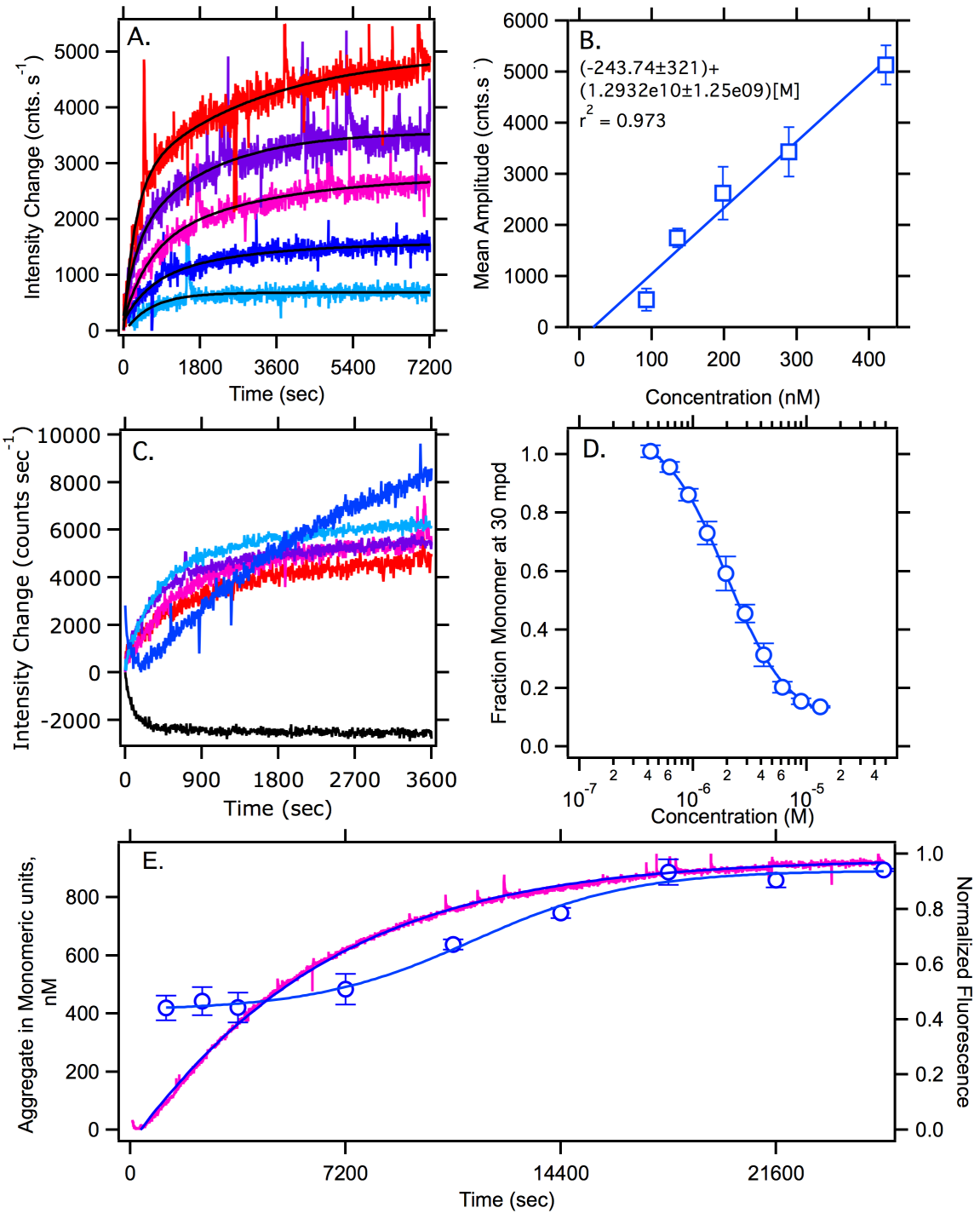


Figure 5.12. Tryptophan Fluorescence and Sedimentation Velocity Kinetics Show uOmpLa Aggregates at Nanomolar Concentrations. Fluorescence (excitation 295 nm, emission 341 nm) and SV kinetics of uOmpLa aggregation in 20 mM Tris-HCl and 150

mM NaCl at pH 8.0, collected at 25 °C. A., uOmpLa kinetics in 80 mM urea depends on protein concentration. Each colored trace is the average of four experiments and black lines are fits to double-exponential functions. Cyan, 92 nM, blue, 135 nM, magenta, 198 nM, purple, 289 nM, and red, 423 nM. B., amplitudes recovered from the fits in A., plotted as a function of uOmpLa concentration. Blue line, extrapolation to zero amplitude gives an intercept of 18.9 ± 25.5 nM. C., Aggregation kinetics (n = 2) of 198 nM uOmpLa depends on urea concentration. Red, 0.08 M, magenta, 0.50 M, purple, 1.0 M, cyan, 2.0 M, blue, 3.0 M, and black, 4.0 M urea. At 4.0 M urea, fluorescence no longer increases over time, suggesting that uOmpLa is monomeric above 4.0 M urea. D., The fraction monomeric uOmpLa measured by SV at 30 minutes post-dilution in 4.0 M urea, plotted versus uOmpLa concentration. The leftmost point, 198 nM in 4.0 M urea, corresponds to the solid black trace in C. Solid blue, a Hill function used to fit the SV data (Danoff and Fleming, 2011). The transition midpoint of the Hill trace is 2.0 ± 0.01 μM. E., Overlay of SV and fluorescence kinetics of 1.33 μM OmpLa in 4.0 M urea. Blue circles, SV data, magenta, fluorescence data. Blue lines are fits to sigmoid functions. Note that tryptophan fluorescence lags behind aggregation at the earliest time points and outpaces it at intermediate times. Finally, both types of kinetics saturate over 24 hours.

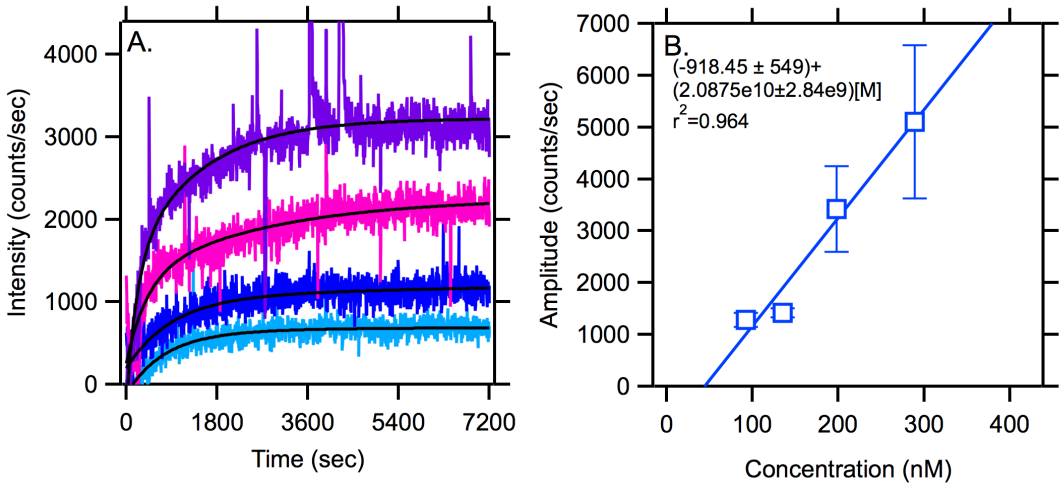


Figure 5.13. BamA Aggregates at a Concentration of 44.0 ± 37.4 nM. A-B., Tryptophan fluorescence kinetics for BamA, collected and analyzed using methods described for figure 5.12A and 5.12B, except that 2 mM TCEP was included in the samples.

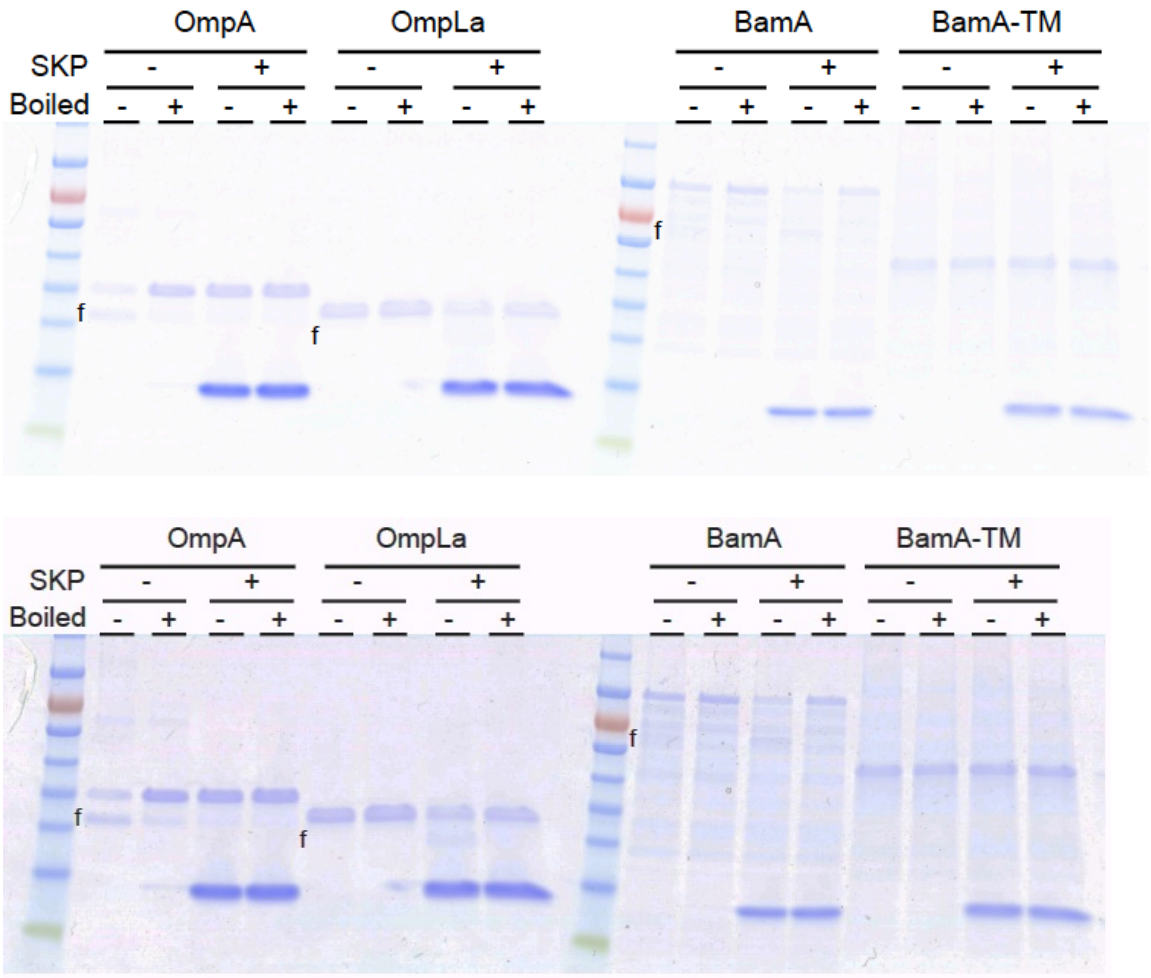


Figure 5.14. Skp Chaperone Activity is Not Correlated with Antiaggregation Activity. Top row, unmodified images, lower row, contrast enhanced by 50%. The listed uOMPs in 6 M urea were rapidly diluted under the same conditions as the AA assays used in SV. At 30 minutes post-dilution, instead of performing SV, uOMPs with or without Skp were added to diC11PC LUVs. The "f" indicates the expected position of the natively folded conformation. Under these conditions at ambient temperature, Skp inhibits the folding of OmpA, and has a limited effect on OmpLa folding. BamA-TM has never displayed a reproducible gel-shift.

Concluding Remarks

In chapter 1, I presented three questions about the structure and function of Skp. These questions were about how uOMPs fit inside Skp, whether Skp is a trimer, and whether Skp is a holding chaperone.

The "cavity paradox" (figure 1.7) was partly resolved by SANS of Skp-uOMP complexes (chapter 2). Different uOMPs shift their centers-of-mass relative to Skp and their parts can protrude from the Skp trimer cavity. Also, Skp trimers expand and contract. Similar adaptation to diverse substrates is seen in multidrug binding proteins (Schumacher and Brennan, 2002), the trigger factor chaperone (Saio, Guan et al., 2014), and Spy (Horowitz, Salmon et al., 2016). The pattern among these systems is that multiple structures are needed to describe them.

Skp is not a constitutive trimer (chapter 4). This implies that simple models of Skp-uOMP binding are not accurate. Because these models are inaccurate, it means that the energies reported for Skp-uOMP binding (Qu, Mayer et al., 2007, Moon, Zaccai et al., 2013) are not reliable. Estimating accurate binding energies is relevant in the context of the thermodynamic hypothesis for OMP sorting (Moon, Zaccai et al., 2013). Thus new models are needed to fit available fluorescence data.

Also in chapter 4, I offered a correction (chapter 4, preface and text) to my interpretation of "flat" linkage plots of salt activity versus trimerization. Such flat linkage plots can result from many mechanisms. For example, there may be simultaneous uptake and release of ions or short-range electrostatic interactions that salts do not screen. Only a few long-range electrostatic mechanisms were ruled out by those negative results.

Skp chaperone activity (CA) for uOMPs is inconsistent, but it operates in heterogeneous conditions that favor Skp misfolding (figure 3.4). These conditions, with 1 M urea, may be thought of as a kind of "stress test" for CA. When subjected to denaturing conditions, Skp worked for the smallest substrate tested (uOmpW) and for a limited amount of time (20 minutes, figure 3.3). This leads me to ask, how well are stress-responsive chaperones able to function *in vivo*, given that they are activated in conditions that favor their unfolding?

Skp antiaggregation activity (AA) does not completely prevent uOMP aggregation at micromolar protein concentrations (figure 5.9). Therefore, the aggregation rate (AR) of uOMPs competes effectively with Skp-uOMP binding rates at these concentrations. This binding rate occurs in milliseconds at the lower concentration of 300 nM (Wu, Ge et al., 2011, Lyu, Shao et al.). If uOMP ARs are this fast at the higher micromolar concentrations then they are close to the fastest uOMP folding rates *in vitro* (Gessmann, Chung et al., 2014). Thus, the AR likely explains the differences in uOMP folding efficiencies (figure 1.3) observed by Burgess et al. (2008). In addition, a slow AR for uOmpA (chapter 5) shows that some "soluble" uOMPs do not require Skp AA (figure 5.9). This raises the question, can "soluble" uOMPs act as competitive inhibitors of Skp AA for "insoluble" uOMPs?

In chapter 1, I explained that holding activity requires many assumptions about kinetic constants (section 1.2 and figure 1.5). The basic assumptions were that uOMPs bound the membrane and folding was not rate-limiting. Another assumption was a rapid release of uOMPs to the membrane from the chaperone (section 1.2 and figure 1.5). However, this requires a chaperone-uOMP-membrane complex. Thus, adding a

membrane to the Skp-uOMP system makes holding activity challenging to observe. This may partly explain why Skp CA is so dependent on membrane components (chapters 3 and 5). To explain the Skp CA, the next step should be to simulate and measure rate constants for a system like that depicted in figure 1.5.

To conclude, I summarize a new model for Skp function in *E. coli* proteostasis (figure 6.1A-B). Under conditions of low envelope stress (figure 6.1A), most Skp is located near the inner membrane (IM) (De Cock, Schäfer et al., 1999). In low stress, Skp is weakly expressed and partly monomeric (De Cock, Schäfer et al., 1999, Sandlin, Zaccai et al., 2015). Also, some Skp trimers interact with a portion of nascent chains at the translocon, while others bind disordered protein (including uOMPs) throughout the periplasm (Harms, Koningstein et al., 2001, Ieva, Tian et al., 2011, Moon, Zaccai et al., 2013, De, Jeong et al., 2014). Skp binding and crowding effects deplete an unknown fraction of Skp monomers (figure 6.1, "(f_{mon}?)"). In this basal growth condition, most of the uOMPs traversing the periplasm diffuse over the surface of PPI-D and/or SurA to the BamABCDE foldase (Gessmann et al. 2014, Plummer & Fleming 2015, Thoma et al. 2015, Wang, Wang et al., 2016). This situation (figure 6.1A) is that observed in rich media, where Skp is nonessential (Sklar, Wu et al., 2007).

Alternatively, environmental factors may inhibit OMP folding rates (6.1B). In this case, sigmaE regulation increases the population of trimeric Skp and other chaperones (Rhodius, Suh et al., 2006, Sandlin, Zaccai et al., 2015). However, if uOMP concentrations saturate the chaperones (PPI-D, SurA, FkpA, and Spy), Skp acts as a "uOMP buffer" absorbing free uOMPs with powerful AA and nanomolar binding affinity. The resulting free uOMP concentrations are kept below the ~10 nM threshold (figure

5.11 and 5.12). uOMPs accumulate in this Skp "kinetic storage tank", unless a ternary complex is formed with DegP. Formation of a complex with DegP results in uOMP degradation (Wu, Ge et al., 2011, Hansen and Hilgenfeld, 2013).

This new model, based on *in vitro* experiments, presents a view that Skp has a speciality: buffering concentrations for uOMPs that are not bound to the other chaperones. The new model predicts that Skp should increase the survival of *E. coli* under conditions where (1) aggregation is favored, (2) aggregation-prone OMP expression is high, and (3) OMP folding rates are slowed. Satisfying these three conditions is challenging in the laboratory, which may explain why Skp deletion is usually phenotypically silent (Sklar, Wu et al., 2007). Perhaps these conditions are satisfied more frequently in nature, where fluctuating solution conditions are common.

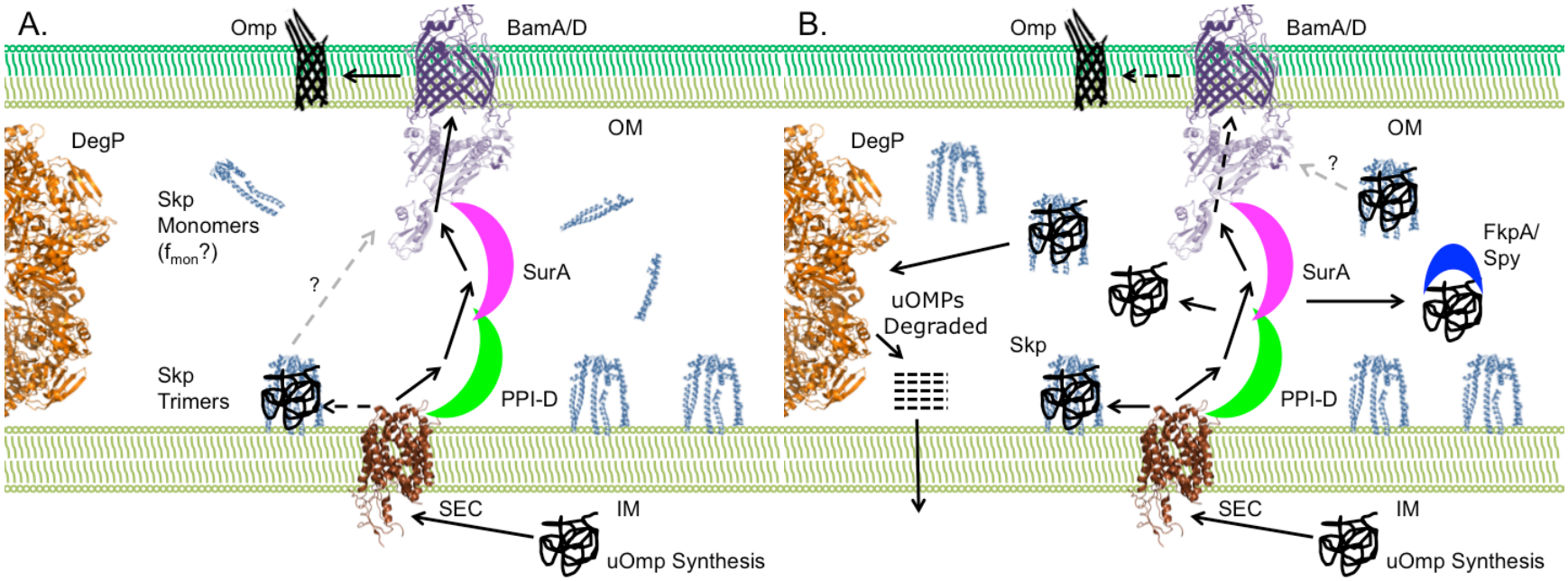


Figure 6.1 A New Model for the Function of Seventeen Kilodalton Protein in *E. coli* Proteostasis. A., in rich media, Skp is weakly expressed and OMPs fold fast. Some fraction of Skp (f_{mon}?) is monomeric, while most is trimeric at the IM. uOMPs diffuse from the translocon (SEC) over PPI-D and SurA to Bam core subunits (BamA/D) that catalyze folding. Skp acts to correct transport errors (dotted arrows). B., when folding is slowed, Skp, FkpA and Spy are upregulated. uOMPs have enough time to dissociate from SurA/PPI-D/FkpA/Spy. Skp trimers in the periplasm act as a buffer and regulator of free uOMPs, keeping their concentration below ~10 nM. Folding from Skp rarely occurs (grey arrow) and many bound uOMPs are degraded by DegP. See main text for references.

INTENDED TO BE BLANK

References

Ades, S. E. (2004). "Control of the alternative sigma factor sigmaE in Escherichia coli." *Curr Opin Microbiol* **7**(2): 157-162.

Aguado, A., J. A. Fernandez-Higuero, F. Moro and A. Muga (2015). "Chaperone-assisted protein aggregate reactivation: Different solutions for the same problem." *Arch Biochem Biophys* **580**: 121-134.

Alekshun, M. N. and S. B. Levy (2006). "Commensals upon us." *Biochem Pharmacol* **71**(7): 893-900.

Arike, L., K. Valgepea, L. Peil, R. Nahku, K. Adamberg and R. Vilu (2012). "Comparison and applications of label-free absolute proteome quantification methods on Escherichia coli." *J Proteomics* **75**(17): 5437-5448.

Balch, W. E., R. I. Morimoto, A. Dillin and J. W. Kelly (2008). "Adapting proteostasis for disease intervention." *Science* **319**(5865): 916-919.

Beaudry, P., H. U. Petersen, M. Grunberg-Manago and B. Jacrot (1976). "A neutron study of the 30 S-ribosome subunit and of the 30 S-IF3 complex." *Biochem Biophys Res Commun* **72**(2): 391-397.

Bednarska, N. G., J. Schymkowitz, F. Rousseau and J. Van Eldere (2013). "Protein aggregation in bacteria: the thin boundary between functionality and toxicity." *Microbiology* **159**(Pt 9): 1795-1806.

Berrow, N. S., D. Alderton, S. Sainsbury, J. Nettleship, R. Assenberg, N. Rahman, D. I. Stuart and R. J. Owens (2007). "A versatile ligation-independent cloning method suitable for high-throughput expression screening applications." *Nucleic Acids Res* **35**(6): e45.

Beverly, K. N., M. R. Sawaya, E. Schmid and C. M. Koehler (2008). "The Tim8-Tim13 complex has multiple substrate binding sites and binds cooperatively to Tim23." *J Mol Biol* **382**(5): 1144-1156.

Bitto, E. and D. McKay (2003). "The periplasmic molecular chaperone protein SurA binds a peptide motif that is characteristic of integral outer membrane proteins." *The Journal of Biological Chemistry* **278**: 49316-49322.

Bratanov, D., T. Balandin, E. Round, V. Shevchenko, I. Gushchin, V. Polovinkin, V. Borshchevskiy and V. Gordeliy (2015). "An Approach to Heterologous Expression of Membrane Proteins. The Case of Bacteriorhodopsin." *PLoS ONE* **10**(6): e0128390.

Bulieris, P. V., S. Behrens, O. Holst and J. H. Kleinschmidt (2003). "Folding and insertion of the outer membrane protein OmpA is assisted by the chaperone Skp and by lipopolysaccharide." *J Biol Chem* **278**(11): 9092-9099.

Burgess, N. K., T. P. Dao, A. M. Stanley and K. G. Fleming (2008). "Beta-barrel proteins that reside in the Escherichia coli outer membrane in vivo demonstrate varied folding behavior in vitro." *J Biol Chem* **283**(39): 26748-26758.

Burmann, B. M. and S. Hiller (2012). "Solution NMR studies of membrane-protein-chaperone complexes." *Chimia (Aarau)* **66**(10): 759-763.

Burmann, B. M., D. A. Holdbrook, M. Callon, P. J. Bond and S. Hiller (2015). "Revisiting the interaction between the chaperone Skp and lipopolysaccharide." *Biophys J* **108**(6): 1516-1526.

Burmann, B. M., C. Wang and S. Hiller (2013). "Conformation and dynamics of the periplasmic membrane-protein-chaperone complexes OmpX-Skp and tOmpA-Skp." *Nat Struct Mol Biol* **20**(11): 1265-1272.

Burmann, B. M. H., D.A.; Callon, M.; Bond, P.J.; Hiller, S. (2015). "Revisiting the Interaction between the Chaperone Skp and Lipopolysaccharide." *Biophys J* **108**(6): 1516-1526.

Cai, G.-Z., L. L.-Y. Lee, M. A. Luther and J. C. Lee (1990). "Regulation and quaternary structural changes in rabbit muscle phosphofructokinase." *Biochemistry* **37**: 97-106.

Callon, M., B. M. Burmann and S. Hiller (2014). "Structural mapping of a chaperone-substrate interaction surface." *Angew Chem Int Ed Engl* **53**(20): 5069-5072.

Cavallo, L., J. Kleinjung and F. Fraternali (2003). "POPS: a fast algorithm for solvent accessible surface areas at atomic and residue level." *Nucleic Acids Research* **31**(13): 3364-3366.

Chatton, J. Y. and K. R. Spring (1995). "The sodium concentration of the lateral intercellular spaces of MDCK cells: a microspectrofluorimetric study." *J Membr Biol* **144**(1): 11-19.

Chen, J., H. S. Won, W. Im, H. J. Dyson and C. L. Brooks, 3rd (2005). "Generation of native-like protein structures from limited NMR data, modern force fields and advanced conformational sampling." *J Biomol NMR* **31**(1): 59-64.

Chen, R. and U. Henning (1996). "A periplasmic protein (Skp) of Escherichia coli selectively binds a class of outer membrane proteins." *Mol Microbiol* **19**(6): 1287-1294.

Clark, N., H. Zhang, S. Krueger, H. Lee, R. Ketchem, B. Kerwin and e. al. (2013). "Small-angle neutron scattering study of a monoclonal antibody using free-energy constraints." *The Journal of Physical Chemistry* **117**: 14029-14038.

Cowan, S. W., T. Schirmer, G. Rummel, M. Steiert, R. Ghosh, R. A. Pauptit, J. N. Jansonius and J. P. Rosenbusch (1992). "Crystal structures explain functional properties of two E. coli porins." *Nature* **358**(6389): 727-733.

Curtis, J. E., S. Raghunandan, H. Nanda and S. Kreuger (2012). "SASSIE: A program to study intrinsically disordered biological molecules and macromolecular ensembles using experimental scattering constraints." *Computer Physics Communications* **183**: 382-389.

Danoff, E. J. and K. G. Fleming (2011). "The soluble, periplasmic domain of OmpA folds as an independent unit and displays chaperone activity by reducing the self-association propensity of the unfolded OmpA transmembrane beta-barrel." *Biophys Chem* **159**(1): 194-204.

Danoff, E. J. and K. G. Fleming (2015). "Aqueous, Unfolded OmpA Forms Amyloid-Like Fibrils upon Self-Association." *PLoS One* **10**(7): e0132301.

Das, R. K. and R. V. Pappu (2013). "Conformations of intrinsically disordered proteins are influenced by linear sequence distributions of oppositely charged residues." *Proc Natl Acad Sci U S A* **110**(33): 13392-13397.

De Cock, H., U. Schäfer, M. Potgeter, R. Demel, M. Müller and J. Tommassen (1999). "Affinity of the periplasmic chaperone Skp of Escherichia coli for phospholipids, lipopolysaccharides and non-native outer membrane proteins. Role of Skp in the biogenesis of outer membrane protein." *Eur J Biochem* **259**(1-2): 96-103.

De, D., M. H. Jeong, Y. E. Leem, D. I. Svergun, D. E. Wemmer, J. S. Kang, K. K. Kim and S. H. Kim (2014). "Inhibition of master transcription factors in pluripotent cells induces early stage differentiation." *Proc Natl Acad Sci U S A* **111**(5): 1778-1783.

Denks, K., A. Vogt, I. Sachelaru, N. A. Petriman, R. Kudva and H. G. Koch (2014). "The Sec translocon mediated protein transport in prokaryotes and eukaryotes." *Mol Membr Biol* **31**(2-3): 58-84.

Denoncin, K., J. Schwalm, D. Vertommen, T. J. Silhavy and J. F. Collet (2012). "Dissecting the Escherichia coli periplasmic chaperone network using differential proteomics." *Proteomics* **12**(9): 1391-1401.

Ding, J., C. Yang, X. Niu, Y. Hu and C. Jin (2015). "HdeB chaperone activity is coupled to its intrinsic dynamic properties." *Sci Rep* **5**: 16856.

Doyle, S. M., J. Shorter, M. Zolkiewski, J. R. Hoskins, S. Lindquist and S. Wickner (2007). "Asymmetric deceleration of ClpB or Hsp104 ATPase activity unleashes protein-remodeling activity." *Nat Struct Mol Biol* **14**(2): 114-122.

Ebie Tan, A., N. K. Burgess, D. S. DeAndrade, J. D. Marold and K. G. Fleming (2010). "Self-association of unfolded outer membrane proteins." *Macromol Biosci* **10**(7): 763-767.

Ellis, J. (1987). "Proteins as molecular chaperones." *Nature* **328**(6129): 378-379.

Ellis, R. J. (2006). "Molecular chaperones: assisting assembly in addition to folding." *Trends Biochem Sci* **31**(7): 395-401.

Ellis, R. J. (2013). "Assembly chaperones: a perspective." *Philos Trans R Soc Lond B Biol Sci* **368**(1617): 20110398.

Engelman, D. M. and P. B. Moore (1975). "Determination of quaternary structure by small angle neutron scattering." *Annu Rev Biophys Bioeng* **4**(00): 219-241.

Entzminger, K. C., C. Chang, R. O. Myhre, K. C. McCallum and J. A. Maynard (2012). "The Skp chaperone helps fold soluble proteins in vitro by inhibiting aggregation." *Biochemistry* **51**(24): 4822-4834.

Fleming, K. G. (2015). "A combined kinetic push and thermodynamic pull as driving forces for outer membrane protein sorting and folding in bacteria." *Philos Trans R Soc Lond B Biol Sci* **370**(1679).

Fleming, P. J., N. C. Fitzkee, M. Mezel, R. Srinivasan and G. Rose (2005). "A novel method reveals that solvent water favors polyproline II over β-strand conformation in peptides and unfolded proteins: conditional hydrophobic accessible surface area (CHASA)." *Protein Science* **14**: 111-118.

Ge, X., Z. X. Lyu, Y. Liu, R. Wang, X. S. Zhao, X. Fu and Z. Chang (2014). "Identification of FkpA as a key quality control factor for the biogenesis of outer membrane proteins under heat shock conditions." *J Bacteriol* **196**(3): 672-680.

Generoso, S. F., M. Giustiniano, G. La Regina, S. Bottone, S. Passacantilli, S. Di Maro, H. Cassese, A. Bruno, M. Mallardo, M. Dentice, R. Silvestri, L. Marinelli, D. Sarnataro, S. Bonatti, E. Novellino and M. Stornaiuolo (2015). "Pharmacological folding chaperones act as allosteric ligands of Frizzled4." *Nat Chem Biol* **11**(4): 280-286.

Gessmann, D., Y. H. Chung, E. J. Danoff, A. M. Plummer, C. W. Sandlin, N. R. Zaccai and K. G. Fleming (2014). "Outer membrane beta-barrel protein folding is physically controlled by periplasmic lipid head groups and BamA." *Proc Natl Acad Sci U S A* **111**(16): 5878-5883.

Geyer, R., C. Galanos, O. Westphal and J. Golecki (1979). "A Lipopolysaccharide-Binding Cell-Surface Protein from *Salmonella minnesota*." *Eur J Biochem* **98**: 27-38.

Gill, S. C. and P. H. von Hippel (1989). "Calculation of protein extinction coefficients from amino acid sequence data." *Anal Biochem.* **182**(2): 319-326.

Glatter, O. (1977). "A new method for the evaluation of small-angle scattering data." *Journal of Applied Crystallography* **10**: 415-421.

Glatter, O. and O. Kratky (1982). *Small-angle X-ray scattering*. New York, Academic Press.

Glinka, C., J. Barker, B. Hammouda, S. Krueger, J. Moyer and W. Orts (1998). "The 30 m small-angle neutron scattering instruments at the national institute of standards and technology." *Journal of Applied Crystallography* **31**: 430-445.

Goloubinoff, P., A. A. Gatenby and G. H. Lorimer (1989). "GroE heat-shock proteins promote assembly of foreign prokaryotic ribulose bisphosphate carboxylase oligomers in Escherichia coli." *Nature* **337**(6202): 44-47.

Goodsell, D. S. (1991). "Inside a living cell." *Trends Biochem Sci* **16**(6): 203-206.

Gray, M. J. and U. Jakob (2015). "Oxidative stress protection by polyphosphate--new roles for an old player." *Curr Opin Microbiol* **24**: 1-6.

Gu, Y., H. Li, H. Dong, Y. Zeng, Z. Zhang, N. G. Paterson, P. J. Stansfeld, Z. Wang, Y. Zhang, W. Wang and C. Dong (2016). "Structural basis of outer membrane protein insertion by the Bam complex." *Nature* **531**(7592): 64-69.

Guiner, A. and G. Fournet (1955). *Small-angle scattering of X-rays*. New York, John Wiley & Sons, Inc.

Gupta, A. J., S. Haldar, G. Milicic, F. U. Hartl and M. Hayer-Hartl (2014). "Active cage mechanism of chaperonin-assisted protein folding demonstrated at single-molecule level." *J Mol Biol* **426**(15): 2739-2754.

Hamer, W. J. and Y. Wu (1972). "Osmotic Coefficients and Mean Activity Coefficients of Univalent Electrolytes in Water at 25 C." *Journal of Physical and Chemical Reference Data* **1**.

Hansen, G. and R. Hilgenfeld (2013). "Architecture and regulation of HtrA-family proteins involved in protein quality control and stress response." *Cell Mol Life Sci* **70**(5): 761-775.

Hansen, S. (1990). "Calculation of small-angle scattering profiles using Monte Carlo simulation." *Journal of Applied Crystallography* **23**: 344-346.

Harms, N., G. Koningstein, W. Dontje, M. Muller, B. Oudega, J. Luirink and H. de Cock (2001). "The early interaction of the outer membrane protein phoe with the periplasmic chaperone Skp occurs at the cytoplasmic membrane." *J Biol Chem* **276**(22): 18804-18811.

Haslbeck, M. and E. Vierling (2015). "A first line of stress defense: small heat shock proteins and their function in protein homeostasis." *J Mol Biol* **427**(7): 1537-1548.

Hayes, D. B., T. Laue, J. Philo, T. Hurton, A. Wright, G. Deubler and B. Bashir (1995). SEDNTERP. Biomolecular Interaction Technologies Center, University of New Hampshire.

Heidorn, D. and J. Trewhella (1988). "Comparison of the crystal and solution structures of calmodulin and troponin C." *Biochemistry* **27**: 909-915.

Helbig, S., S. I. Patzer, C. Schiene-Fischer, K. Zeth and V. Braun (2011). "Activation of colicin M by the FkpA prolyl cis-trans isomerase/chaperone." *J Biol Chem* **286**(8): 6280-6290.

Hiller, S. and G. Wagner (2009). "The role of solution NMR in the structure determinations of VDAC-1 and other membrane proteins." *Curr Opin Struct Biol* **19**(4): 396-401.

Hoffmann, A., B. Bukau and G. Kramer (2010). "Structure and function of the molecular chaperone Trigger Factor." *Biochim Biophys Acta* **1803**(6): 650-661.

Holck, A. and K. Kleppe (1988). "Cloning and sequencing of the gene for the DNA-binding 17K protein of Escherichia coli." *Gene* **67**(1): 117-124.

Hong, H., D. R. Patel, L. K. Tamm and B. van den Berg (2006). "The outer membrane protein OmpW forms an eight-stranded beta-barrel with a hydrophobic channel." *J Biol Chem* **281**(11): 7568-7577.

Ibel, K. and H. B. Stuhrmann (1975). "Comparison of neutron and X-ray scattering of dilute myoglobin solutions." *Journal of Molecular Biology* **93**: 255-265.

Ieva, R., P. Tian, J. H. Peterson and H. D. Bernstein (2011). "Sequential and spatially restricted interactions of assembly factors with an autotransporter beta domain." *Proc Natl Acad Sci U S A* **108**(31): E383-391.

Ishimoto, T., K. Fujiwara, T. Niwa and H. Taguchi (2014). "Conversion of a chaperonin GroEL-independent protein into an obligate substrate." *J Biol Chem* **289**(46): 32073-32080.

Jacques, D. A. and J. Trewhella (2010). "Small-angle scattering for structural biology--expanding the frontier while avoiding the pitfalls." *Protein Sci* **19**(4): 642-657.

Jarchow, S., C. Luck, A. Gorg and A. Skerra (2008). "Identification of potential substrate proteins for the periplasmic Escherichia coli chaperone Skp." *Proteomics* **8**(23-24): 4987-4994.

Jaroslawski, S., K. Duquesne, J. N. Sturgis and S. Scheuring (2009). "High-resolution architecture of the outer membrane of the Gram-negative bacteria Roseobacter denitrificans." *Mol Microbiol* **74**(5): 1211-1222.

Johnson, M. L., J. J. Correia, D. A. Yphantis and H. R. Halvorson (1981). "Analysis of data from the analytical ultracentrifuge by nonlinear least-squares techniques." *Biophys J* **36**(3): 575-588.

Kim, D. Y. (2015). "Two stress sensor proteins for the expression of sigmaE regulon: DegS and RseB." *J Microbiol* **53**(5): 306-310.

Kleinschmidt, J. H. (2015). "Folding of beta-barrel membrane proteins in lipid bilayers - Unassisted and assisted folding and insertion." *Biochim Biophys Acta* **1848**(9): 1927-1943.

Kline, S. (2006). "Reduction and analysis of SANS and USANS data using IGOR Pro." *Journal of Applied Crystallography* **39**: 895-900.

Koli, P., S. Sudan, D. Fitzgerald, S. Adhya and S. Kar (2011). "Conversion of commensal Escherichia coli K-12 to an invasive form via expression of a mutant histone-like protein." *MBio* **2**(5).

Korndörfer, I. P., M. K. Dommel and A. Skerra (2004). "Structure of the periplasmic chaperone Skp suggests functional similarity with cytosolic chaperones despite differing architecture." *Nat Struct Mol Biol* **11**(10): 1015-1020.

Krueger, S., I. Gorshkova, J. Brown, J. Hoskins, K. McKenney and F. Schwartz (1998). "Determination of the conformations of the cAMP receptor protein and its T127L, S128A mutant with and without cAMP from small angle neutron scattering measurements." *Journal of Biological Chemistry* **273**: 20001-20006.

Krueger, S., J. H. Shin, J. E. Curtis, K. A. Rubinson and Z. Kelman (2014). "The solution structure of full-length dodecameric MCM by SANS and molecular modeling." *Proteins* **82**(10): 2364-2374.

Krueger, S., J. H. Shin, S. Raghunandan, J. E. Curtis and Z. Kelman (2011). "Atomistic ensemble modeling and small-angle neutron scattering of intrinsically disordered protein complexes: applied to minichromosome maintenance protein." *Biophys J* **101**(12): 2999-3007.

Kubota, H. (2009). "Quality control against misfolded proteins in the cytosol: a network for cell survival." *J Biochem* **146**(5): 609-616.

Laskey, R. A., B. M. Honda, A. D. Mills and J. T. Finch (1978). "Nucleosomes are assembled by an acidic protein which binds histones and transfers them to DNA." *Nature* **275**(5679): 416-420.

Lee, K. K., C. A. Fitch and B. García-Moreno E. (2002). "Distance dependence and salt sensitivity of pairwise, coulombic interactions in a protein." *Protein Science* **11**: 1004-1016.

Leonard-Rivera, M. and R. Misra (2012). "Conserved residues of the putative L6 loop of Escherichia coli BamA play a critical role in the assembly of beta-barrel outer membrane proteins, including that of BamA itself." *J Bacteriol* **194**(17): 4662-4668.

Liu, Y. and J. M. Sturtevant (1995). "Significant discrepancies between van't Hoff and calorimetric enthalpies. II." *Protein Science* **4**: 2559-2561.

Lu, J. and A. Holmgren (2014). "The thioredoxin superfamily in oxidative protein folding." *Antioxid Redox Signal* **21**(3): 457-470.

Lundin, V. F., M. R. Leroux and P. C. Stirling (2010). "Quality control of cytoskeletal proteins in human disease." *Trends Biochem Sci* **35**(5): 288-297.

Lyu, Z. X., Q. Shao, Y. Q. Gao and X. S. Zhao (2012). "Direct observation of the uptake of outer membrane proteins by the periplasmic chaperone Skp." *PLoS One* **7**(9): e46068.

MacKerell, A. D., D. Bashford, M. Bellott, R. L. Dunbrack, J. D. Evanseck, M. J. Field, S. Fischer, J. Gao, H. Guo, S. Ha, D. Joseph-McCarthy, L. Kuchnir, K. Kuczera, F. T. Lau, C. Mattos, S. Michnick, T. Ngo, D. T. Nguyen, B. Prodhom, W. E. Reiher, B. Roux, M. Schlenkrich, J. C. Smith, R. Stote, J. Straub, M. Watanabe, J. Wiorkiewicz-Kuczera, D. Yin and M. Karplus (1998). "All-atom empirical potential for molecular modeling and dynamics studies of proteins." *J Phys Chem B* **102**(18): 3586-3616.

Malmberg, C. G. and A. A. Maryott (1956). "Dielectric Constant of Water from 0 to 100 C." *Journal of Research of the National Bureau of Standards* **56**(1).

Martin-Benito, J., J. Boskovic, P. Gomez-Puertas, J. L. Carrascosa, C. T. Simons, S. A. Lewis, F. Bartolini, N. J. Cowan and J. M. Valpuesta (2002). "Structure of eukaryotic prefoldin and of its complexes with unfolded actin and the cytosolic chaperonin CCT." *Embo j* **21**(23): 6377-6386.

Martin-Benito, J., J. Gomez-Reino, P. C. Stirling, V. F. Lundin, P. Gomez-Puertas, J. Boskovic, P. Chacon, J. J. Fernandez, J. Berenguer, M. R. Leroux and J. M. Valpuesta (2007). "Divergent substrate-binding mechanisms reveal an evolutionary specialization of eukaryotic prefoldin compared to its archaeal counterpart." *Structure* **15**(1): 101-110.

Masuda, T., N. Saito, M. Tomita and Y. Ishihama (2009). "Unbiased quantitation of Escherichia coli membrane proteome using phase transfer surfactants." *Mol Cell Proteomics* **8**(12): 2770-2777.

Matern, Y., B. Barion and S. Behrens-Kneip (2010). "PpiD is a player in the network of periplasmic chaperones in Escherichia coli." *BMC Microbiol* **10**: 251.

Mattoo, R. U. and P. Goloubinoff (2014). "Molecular chaperones are nanomachines that catalytically unfold misfolded and alternatively folded proteins." *Cell Mol Life Sci* **71**(17): 3311-3325.

McMorran, L. M., A. I. Bartlett, G. H. Huysmans, S. E. Radford and D. J. Brockwell (2013). "Dissecting the effects of periplasmic chaperones on the in vitro folding of the outer membrane protein PagP." *J Mol Biol* **425**(17): 3178-3191.

Meng, F., Y. Park and H. Zhou (2001). "Role of proline, glycerol, and heparin as protein folding aids during refolding of rabbit muscle creatine kinase." *Int J Biochem Cell Biol* **33**(7): 701-709.

Millan-Zambrano, G. and S. Chavez (2014). "Nuclear functions of prefoldin." *Open Biol* **4**(7).

Missiakas, D., J. M. Betton and S. Raina (1996). "New components of protein folding in extracytoplasmic compartments of Escherichia coli SurA, FkpA and Skp/OmpH." *Mol Microbiol* **21**(4): 871-884.

Moon, C. P. (2011). *Hydrophobicity and the Thermodynamic Stability of Membrane Proteins*. Molecular Biophysics Doctoral, Johns Hopkins University.

Moon, C. P., N. R. Zaccai, P. J. Fleming, D. Gessmann and K. G. Fleming (2013). "Membrane protein thermodynamic stability may serve as the energy sink for sorting in the periplasm." *Proc Natl Acad Sci U S A* **110**(11): 4285-4290.

Moore, P. (1982). *Small-angle scattering techniques for the study of biological macromolecules and macromolecular aggregates.*, Academic Press.

Nakamura, K. and S. Mizushima (1976). "Effects of heating in dodecyl sulfate solution on the conformation and electrophoretic mobility of isolated major outer membrane proteins from Escherichia coli K-12." *J Biochem* **80**(6): 1411-1422.

Patel, G. J., S. Behrens-Kneip, O. Holst and J. H. Kleinschmidt (2009). "The periplasmic chaperone Skp facilitates targeting, insertion, and folding of OmpA into lipid membranes with a negative membrane surface potential." *Biochemistry* **48**(43): 10235-10245.

Patel, G. J. and J. H. Kleinschmidt (2013). "The lipid bilayer-inserted membrane protein BamA of Escherichia coli facilitates insertion and folding of outer membrane protein A from its complex with Skp." *Biochemistry* **52**(23): 3974-3986.

Pautsch, A. and G. E. Schulz (1998). "Structure of the outer membrane protein A transmembrane domain." *Nat Struct Biol* **5**(11): 1013-1017.

Peng, Y., J. E. Curtis, X. Fang and S. Woodson (2014). "Structural model of an mRNA in complex with the bacterial chaperone Hfq." *Proc Natl Acad Sci U S A* **111**: 17134-17139.

Phillips, J. C., R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R. D. Skeel, L. Kale and K. Schulten (2005). "Scalable molecular dynamics with NAMD." *J Comput Chem* **26**(16): 1781-1802.

Philo, J. S. (2006). "Improved methods for fitting sedimentation coefficient distributions derived by time-derivative techniques." *Anal Biochem* **354**(2): 238-246.

Plummer, A. M. and K. G. Fleming (2015). "BamA Alone Accelerates Outer Membrane Protein Folding In Vitro through a Catalytic Mechanism." *Biochemistry* **54**(39): 6009-6011.

Prabhu, N. V. and K. A. Sharp (2005). "Heat Capacity in Proteins." *Annual Reviews in Physical Chemistry* **56**: 521-548.

Putnam, C. D., M. Hammel, G. L. Hura and J. A. Tainer (2007). "X-ray solution scattering (SAXS) combined with crystallography and computation: defining accurate macromolecular structures, conformations and assemblies in solution." *Q Rev Biophys* **40**(3): 191-285.

Qu, J. (2007). *The interactions of outer membrane proteins with the periplasmic chaperone Skp of E. coli and with LPS*. Naturwissenschaften Doctoral, Universitat Konstanz.

Qu, J., S. Behrens-Kneip, O. Holst and J. H. Kleinschmidt (2009). "Binding regions of outer membrane protein A in complexes with the periplasmic chaperone Skp. A site-directed fluorescence study." *Biochemistry* **48**(22): 4926-4936.

Qu, J., C. Mayer, S. Behrens, O. Holst and J. H. Kleinschmidt (2007). "The trimeric periplasmic chaperone Skp of Escherichia coli forms 1:1 complexes with outer membrane proteins via hydrophobic and electrostatic interactions." *J Mol Biol* **374**(1): 91-105.

Rambo, R. P. and J. A. Tainer (2010). "Bridging the solution divide: comprehensive structural analyses of dynamic RNA, DNA, and protein assemblies by small-angle X-ray scattering." *Curr Opin Struct Biol* **20**(1): 128-137.

Rassam, P., N. A. Copeland, O. Birkholz, C. Toth, M. Chavent, A. L. Duncan, S. J. Cross, N. G. Housden, R. Kaminska, U. Seger, D. M. Quinn, T. J. Garrod, M. S. Sansom, J. Piehler, C. G. Baumann and C. Kleanthous (2015). "Supramolecular assemblies underpin turnover of outer membrane proteins in bacteria." *Nature* **523**(7560): 333-336.

Rhodius, V. A., W. C. Suh, G. Nonaka, J. West and C. A. Gross (2006). "Conserved and variable functions of the sigmaE stress response in related genomes." *PLoS Biol* **4**(1): e2.

Rizzitello, A. E., J. R. Harper and T. J. Silhavy (2001). "Genetic evidence for parallel pathways of chaperone activity in the periplasm of Escherichia coli." *J Bacteriol* **183**(23): 6794-6800.

Robertson, A. D. and K. P. Murphy (1997). "Protein Structure and the Energetics of Protein Stability." *Chemical Reviews* **97**: 1251-1267.

Rollauer, S. E., M. A. Sooreshjani, N. Noinaj and S. K. Buchanan (2015). "Outer membrane protein biogenesis in Gram-negative bacteria." *Philos Trans R Soc Lond B Biol Sci* **370**(1679).

Sachelaru, I., N. A. Petriman, R. Kudva and H. G. Koch (2014). "Dynamic interaction of the sec translocon with the chaperone PpiD." *J Biol Chem* **289**(31): 21706-21715.

Saio, T., X. Guan, P. Rossi, A. Economou and C. G. Kalodimos (2014). "Structural basis for protein antiaggregation activity of the trigger factor chaperone." *Science* **344**(6184): 1250494.

Sandlin, C. W., N. R. Zaccai and K. G. Fleming (2015). "Skp Trimer Formation is Insensitive to Salts in the Physiological Range." *Biochemistry* **54**: 7059−7062.

Sarachan, K., J. Curtis and S. Kreuger (2013). "Small-angle scattering contrast calculator for protein and nucleic acid complexes in solution." *Journal of Applied Crystallography* **46**: 1889-1893.

Schäfer, U., K. Beck and M. Müller (1999). "Skp, a molecular chaperone of gram-negative bacteria, is required for the formation of soluble periplasmic intermediates of outer membrane proteins." *J Biol Chem* **274**(35): 24567-24574.

Schiene-Fischer, C. (2015). "Multidomain Peptidyl Prolyl cis/trans Isomerases." *Biochim Biophys Acta* **1850**(10): 2005-2016.

Schlapschy, M., M. K. Dommel, K. Hadian, M. Fogarasi, I. P. Korndorfer and A. Skerra (2004). "The periplasmic *E. coli* chaperone Skp is a trimer in solution: biophysical and preliminary crystallographic characterization." *Biol Chem* **385**(2): 137-143.

Schulz, G. E. (2002). "The structure of bacterial outer membrane proteins." *Biochim Biophys Acta* **1565**(2): 308-317.

Schwalm, J., T. F. Mahoney, G. R. Soltes and T. J. Silhavy (2013). "Role for Skp in LptD assembly in Escherichia coli." *J Bacteriol* **195**(16): 3734-3742.

Schwede, T., J. Kopp, N. Guex and M. C. Peitsch (2003). "SWISS-MODEL: An automated protein homology-modeling server." *Nucleic Acids Res* **31**(13): 3381-3385.

Semenyuk, A. and D. Svergun (1991). "GNOM-A program package for small-angle scattering data processing." *Journal of Applied Crystallography* **24**: 537-540.

Siegert, R., M. R. Leroux, C. Scheufler, F. U. Hartl and I. Moarefi (2000). "Structure of the molecular chaperone prefoldin: unique interaction of multiple coiled coil tentacles with unfolded proteins." *Cell* **103**(4): 621-632.

Sklar, J. G., T. Wu, D. Kahne and T. J. Silhavy (2007). "Defining the roles of the periplasmic chaperones SurA, Skp, and DegP in Escherichia coli." *Genes Dev* **21**(19): 2473-2484.

Stafford, W. F., 3rd (1992). "Boundary analysis in sedimentation transport experiments: a procedure for obtaining sedimentation coefficient distributions using the time derivative of the concentration profile." *Anal Biochem* **203**(2): 295-301.

Stirling, P. C., S. F. Bakhoum, A. B. Feigl and M. R. Leroux (2006). "Convergent evolution of clamp-like binding sites in diverse chaperones." *Nat Struct Mol Biol* **13**(10): 865-870.

Stock, J. B., B. Rauch and S. Roseman (1977). "Periplasmic space in Salmonella typhimurium and Escherichia coli." *J Biol Chem* **252**(21): 7850-7861.

Stull, F., P. Koldewey, J. R. Humes, S. E. Radford and J. C. Bardwell (2016). "Substrate protein folds while it is bound to the ATP-independent chaperone Spy." *Nat Struct Mol Biol* **23**(1): 53-58.

Tao, H., C. Bausch, C. Richmond, F. R. Blattner and T. Conway (1999). "Functional Genomics: Expression Analysis of *Escherichia coli* Growing on Minimal and Rich Media." *Journal of Bacteriology* **181**(20): 6425-6440.

Tapley, T. L., T. M. Franzmann, S. Chakraborty, U. Jakob and J. C. Bardwell (2010). "Protein refolding by pH-triggered chaperone binding and release." *Proc Natl Acad Sci U S A* **107**(3): 1071-1076.

Teller, D. C. (1973). "Characterization of proteins by sedimentation equilibrium in the analytical ultracentrifuge." *Methods Enzymol* **27**: 346-441.

Thoma, J., B. M. Burmann, S. Hiller and D. J. Muller (2015). "Impact of holdase chaperones Skp and SurA on the folding of beta-barrel outer-membrane proteins." *Nat Struct Mol Biol* **22**(10): 795-802.

Tuttle, M. D., G. Comellas, A. J. Nieuwkoop, D. J. Covell, D. A. Berthold, K. D. Kloepper, J. M. Courtney, J. K. Kim, A. M. Barclay, A. Kendall, W. Wan, G. Stubbs, C. D. Schwieters, V. M. Lee, J. M. George and C. M. Rienstra (2016). "Solid-state NMR structure of a pathogenic fibril of full-length human alpha-synuclein." *Nat Struct Mol Biol* **23**(5): 409-415.

Vergnolle, M. A., C. Baud, A. P. Golovanov, F. Alcock, P. Luciano, L. Y. Lian and K. Tokatlidis (2005). "Distinct domains of small Tims involved in subunit interaction and substrate recognition." *J Mol Biol* **351**(4): 839-849.

Vogt, J. and G. E. Schulz (1999). "The structure of the outer membrane protein OmpX from Escherichia coli reveals possible mechanisms of virulence." *Structure* **7**(10): 1301-1309.

Voulhoux, R., M. P. Bos, J. Geurtsen, M. Mols and J. Tommassen (2003). "Role of a highly conserved bacterial protein in outer membrane protein assembly." *Science* **299**(5604): 262-265.

Walter, S. and J. Buchner (2002). "Molecular chaperones--cellular machines for protein folding." *Angew Chem Int Ed Engl* **41**(7): 1098-1113.

Walton, T. A., C. M. Sandoval, C. A. Fowler, A. Pardi and M. C. Sousa (2009). "The cavity-chaperone Skp protects its substrate from aggregation but allows independent folding of substrate domains." *Proc Natl Acad Sci U S A* **106**(6): 1772-1777.

Walton, T. A. and M. C. Sousa (2004). "Crystal structure of Skp, a prefoldin-like chaperone that protects soluble and membrane proteins from aggregation." *Mol Cell* **15**(3): 367-374.

Webb, C. T., M. A. Gorman, M. Lazarou, M. T. Ryan and J. M. Gulbis (2006). "Crystal structure of the mitochondrial chaperone TIM9.10 reveals a six-bladed alpha-propeller." *Mol Cell* **21**(1): 123-133.

Webb, C. T., E. Heinz and T. Lithgow (2012). "Evolution of the beta-barrel assembly machinery." *Trends Microbiol* **20**(12): 612-620.

Whitman, W. B., D. C. Coleman and W. J. Wiebe (1998). "Prokaryotes: the unseen majority." *Proc Natl Acad Sci U S A* **95**(12): 6578-6583.

Whitten, A., S. Cai and J. Trewhella (2008). "MULCh: Modules for the analysis of small-angle neutron contrast variation data from biomolecular assemblies." *Journal of Applied Crystallography* **41**: 222-226.

Whitten, A. E., D. A. Jacques, B. Hammouda, T. Hanley, G. F. King, J. M. Guss, J. Trewhella and D. B. Langley (2007). "The structure of the KinA-Sda complex suggests an allosteric mechanism of histidine kinase inhibition." *J Mol Biol* **368**(2): 407-420.

Wilkins, D. K., S. B. Grimshaw, V. Receveur, C. M. Dobson, J. A. Jones and L. J. Smith (1999). "Hydrodynamic radii of native and denatured proteins measured by pulse field gradient NMR techniques." *Biochemistry* **38**(50): 16424-16431.

Wu, S., X. Ge, Z. Lv, Z. Zhi, Z. Chang and X. S. Zhao (2011). "Interaction between bacterial outer membrane proteins and periplasmic quality control factors: a kinetic partitioning mechanism." *Biochem J* **438**(3): 505-511.

Xie, Y. (2010). "Structure, assembly and homeostatic regulation of the 26S proteasome." *J Mol Cell Biol* **2**(6): 308-317.

Yohannes, E., D. M. Barnhart and J. L. Slonczewski (2004). "pH-dependent catabolic protein expression during anaerobic growth of Escherichia coli K-12." *J Bacteriol* **186**(1): 192-199.

Young, J. C., V. R. Agashe, K. Siegers and F. U. Hartl (2004). "Pathways of chaperone-mediated protein folding in the cytosol." *Nat Rev Mol Cell Biol* **5**(10): 781-791.

Yphantis, D. A. (1964). "EQUILIBRIUM ULTRACENTRIFUGATION OF DILUTE SOLUTIONS." *Biochemistry* **3**: 297-317.

Zaccai, G. (2012). "Straight lines of neutron scattering in biology: a review of basic controls in SANS and EINS." *Eur Biophys J* **41**(10): 781-787.

Zaccai, N. R., C. W. Sandlin, J. T. Hoopes, J. E. Curtis, P. J. Fleming, K. G. Fleming and S. Krueger (2015). "Deuterium Labeling Together with Contrast Variation Small-Angle Neutron Scattering Suggests How Skp Captures and Releases Unfolded Outer Membrane Proteins." *Methods in Enzymology* **Under Review**.

Zaccai, N. R., C. W. Sandlin, J. T. Hoopes, J. E. Curtis, P. J. Fleming, K. G. Fleming and S. Krueger (2016). "Deuterium Labeling Together with Contrast Variation Small-Angle Neutron Scattering Suggests How Skp Captures and Releases Unfolded Outer Membrane Proteins." *Methods Enzymol* **566**: 159-210.

Appendix I: Commentary on the Need to Improve to the Classic Definition of a Molecular Chaperone

While researching chaperones, I learned that the classic definition of a molecular chaperone is obsolete. An updated definition was beyond the scope of this work (and the patience of my committee). Therefore I will sketch out my findings here. Briefly, the definition of a "molecular chaperone" (Laskey, Honda et al., 1978, Ellis, 1987, Ellis, 2006, Ellis, 2013), invented by Ellis, contains two requirements that are contradicted by recent research advances. First, the definition says that a candidate for chaperone status must be a protein. Numerous examples of non-protein chaperones have recently been identified (Meng, Park et al., 2001, Generoso, Giustiniano et al., 2015, Gray and Jakob, 2015). Treating these non-protein chaperone activities (CAs) differently may not be appropriate. This "protein requirement" may be limiting our search for CAs: Conspicuously, I was not able to find any studies showing RNA CA, despite the powerful CA of polyphosphate (Gray and Jakob, 2015), which is chemically similar to RNA. Second, the old definition says that a putative chaperone is not allowed to be part of the structure of its prospective client. However, HSP104 was recently found to bind the native state of its clients (Doyle, Shorter et al., 2007). Also, the BamA protein is a membrane foldase that can accelerate its own folding (Leonard-Rivera and Misra, 2012). Likewise, OmpA's periplasmic domain has CA (Danoff and Fleming, 2011). A new definition could account for these situations. Definitions are essential to organize reasoning. A better definition might help us identify the structural basis of CA and prevent proteins like Skp from being misclassified. I suggest the definition should be exact and depend upon experimental observations of kinetic CA.

Appendix II: A Python Script Used to Calculate ξ for uOMPs (section 5.2)

```
# coding=utf-8
# This script takes as an input a string of amino acids and calculates the
# hydrophobic sequence segregation.
# This parameter, xi, is similar to the parameter Kappa invented by Pappu Lab.
# xi is calculated using a parameter sigmaH, which is the hydrophobic asymmetry
# parameter.
# sigmaH is analogous to the charge asymmetry used to calculate Kappa.

import math

# dictionary of Preston Moon's water to lipid transfer free energy data, shifted to the
value of lysine, -0.01 kcal,
# and made positive
DDG = {'A': 5.40,
        'C': 4.91,
        'D': 2.45,              #<4.61RT
        'E': 3.76,
        'F': 7.60,
        'G': 3.68,
        'H': 0.64,              #<4.61RT
        'I': 6.96,
        'K': 0.01,              #<4.61RT
        'L': 7.21,
        'M': 6.16,
        'N': 1.93,              #<4.61RT
        'P': 6.92,
        'Q': 2.39,              #<4.61RT
        'R': 1.69,              #<4.61RT
        'S': 3.57,
        'T': 3.62,
        'V': 6.18,
        'W': 5.78,
        'Y': 4.31}

# Dictionary DDG reversed so that the energies are the indexed item and the amino
acids the thing pointed to

inv_DDG = {v: k for k, v in DDG.items()}

# Functions section

#the sigmaH function calculates the sigmaH parameter for an input sequence, which
may be the entire amino acid sequence, or a segment called a "blob"
def sigmaH(hydro):
        a=hydro[0]
```

```
        b=hydro[1]
        sigma=b**2/(a+b)
        return sigma

# meanDDG function finds the sum of the p = (DDGs<RT)/N and the sum of the
significant h=DDGs/N
def meanDDG(seq):
        n=len(seq)
        p=0
        h=0
        for i in seq:
                #print DDG[i]
                if DDG[i]>2.73:                 #Stores all DDGs for residues that are
100-fold more likely to be buried than lysine
                        h=h+DDG[i]
                        #print 'h=',h
                else:
                        p=p+DDG[i]              #Stores the sum of all the DDGs in the
sequence for the rest of the residues not added to h
                        #print 'p=',p
        H=h/n
        P=p/n
        #print [P,H]
        return [P,H]

# This is the main body of the program. It takes an input sequence, calculates the
overall hydrophobic asymmetry, then walks through the sequence twice, first in steps
of 5 and then in steps of 6. It records the local asymmetry in vectors called blobs5 and
blobs6. Next, it calculates deltaH5 and deltaH6 as the variance of the local asymmetry
with respect to the overall asymmetry. It goes on to calculate the maximum possible
deltaH, by arranging the DDGs from smallest to largest value. Finally, it calculates xi,
as the ratio of deltaH to the maximum possible deltaH.

f=open('sequence.txt','rb')
sequence=f.read()                       #Assigns the first line of f to sequence. This is
the amino acid sequence.
f.close()

N = len(sequence)                       #Number of amino acids
Nblobs5=N-4
Nblobs6=N-5

#-------------------------------------------------------------------------------
--
#This section calculates the deltaH for blobs of length 5 and length 6
HasymmOV = sigmaH(meanDDG(sequence))

#This part walks through the blobs of size 5 and records the local sigmaHs.
blobs5=[]
s =[]
```

```
i=0
while (i+1)<Nblobs5:
        s=sequence[i:i+5]
        #print s
        blobs5.append(sigmaH(meanDDG(s)))
        i=i+1
#This part walks through the blobs of size 6 and records the local sigmaHs.
blobs6=[]
s=[]
i=0
while (i+1)<Nblobs6:
        s=sequence[i:i+6]
        blobs6.append(sigmaH(meanDDG(s)))
        i=i+1

#Walks through Nblobs5 and calculates the variance with respect to HasymmOV.
SSR=0
for blob in blobs5:
        SSR=SSR+(blob-HasymmOV)**2

deltaH5=SSR/Nblobs5                #Calculates deltaH for sigmas in blobs5 relative
to HasymmOV

#Walks through Nblobs6 and calculates the variance with respect to HasymmOV.
SSR=0
for blob in blobs6:
        SSR=SSR+(blob-HasymmOV)**2

deltaH6=SSR/Nblobs6                #Calculates deltaH for sigmas in blobs6 relative
to HasymmOV

#-------------------------------------------------------------------------------
—

#This section calculates the maximum deltaH possible, deltaHmax, with all DDGs
arranged in order
#along a sequence.
DDGseq=[]                  #Holds the transfer free energies for each amino acid in
sequence.

for j in sequence:
        DDGseq.append(DDG[j])                    # Creates a vector of the DDGs
from the amino acids.

DDGseq=sorted(DDGseq)                            #Sorts the DDGs in DDGseq
from lowest to highest
                                                 #(creates most segregated
possible sequence)
maxseq = []                                      #Holds the maximally segregated
sequence of amino acids
```

```
for j in DDGseq:
        maxseq.append(inv_DDG[j])          #calls inv_DDG to get the amino acids
for each DDG

HasymmOVmax = sigmaH(meanDDG(maxseq))     #calculates the maximum
possible overall asymmetry.

#This part walks through the blobs in steps of 5 and records the local sigmaHs.
blobs5=[]
s =[]
i=0
while (i+1)<Nblobs5:
        s=maxseq[i:i+5]
        blobs5.append(sigmaH(meanDDG(s)))
        i=i+1
#This part walks through the blobs in steps of 6 and records the local sigmaHs.
blobs6=[]
s=[]
i=0
while (i+1)<Nblobs6:
        s=maxseq[i:i+6]
        blobs6.append(sigmaH(meanDDG(s)))
        i=i+1

#Walks through blobs5 and calculates the variance with respect to HasymmOVmax.

SSR=0
for blob in blobs5:
        SSR=SSR+(blob-HasymmOVmax)**2

deltaH5max=SSR/Nblobs5              #Calculates deltaH for sigmas in blobs5 relative
to HasymmOVmax

#Walks through blobs6 and calculates the variance with respect to HasymmOVmax.

SSR=0
for blob in blobs6:
        SSR=SSR+(blob-HasymmOVmax)**2

deltaH6max=SSR/Nblobs6              #Calculates deltaH for sigmas in blobs6 relative
to HasymmOVmax


# This final section calculates xi, the ratio of the average delta's for the specific
sequence to the average
#delta's for the maximally segregated sequence.

print 'The input sequence was'
print sequence
```

```
print
print 'The maximally segregated sequence was sorted as'
print maxseq
print
print 'The number of amino acids is: ', N
print 'Overall Asymmetry:', HasymmOV
print 'The blobs of length 5 were: ', blobs5
print 'The blobs of length 6 were: ', blobs6
print 'deltaH:', (deltaH5+deltaH6)/2
print 'deltaHmax:', (deltaH5max+deltaH6max)/2
print
print 'XI = ', (deltaH5+deltaH6)/(deltaH5max+deltaH6max)
```

Curriculum Vitae

CLIFFORD W. SANDLIN

940-297-5628 / csandli1@jhu.edu / 4414 Laplata Ave Apt L, Baltimore, MD 21211

OBJECTIVE

I want to answer interesting questions about the molecular factors that underlie maintenance of homeostasis and the role of cell membranes in both homeostasis and symbiosis.

FUTURE RESEARCH INTERESTS

I seek to understand how cell membranes function in the maintenance of homeostasis and symbiosis. To further this understanding, I am interested in using quantitative methodologies and experiments to study how macromolecules and membranes interact and function in these processes.

EDUCATION

Expected Graduation, Summer 2016

Ph.D. Candidate, Program in Molecular Biophysics, The Johns Hopkins University

2011, Bachelor of Science, Mathematics and Biology, University of North Texas

2001, Bachelor of Science, Business Administration, University of Texas at Dallas

1995, Plano Senior High School

EMPLOYMENT

2011-Present, Graduate Student, Program in Molecular Biophysics, Johns Hopkins University, Baltimore, MD

2014, Fall, Teaching Assistant, Computational Biology, Johns Hopkins University, Baltimore, MD

2013, Fall, Teaching Assistant, Cellular and Molecular Physiology, Johns Hopkins University, Baltimore, MD

2010, Summer Intern, NSF Research Experiences for Undergraduates, University of North Texas, Denton, TX

2009, Summer Intern, NSF Research Experiences for Undergraduates, University of Arizona, Tucson AZ

2008, Research Assistant, Benthic Invertebrate Quantification, James Kennedy Lab, Denton, TX

2008, Summer Intern, Student Conservation Association, Indiana Dunes National Lakeshore, IN

2007, Research Assistant (Volunteer), Toxicology, Aaron Roberts Lab, Denton, TX

2003-2007, Book Binder, Fedex Kinkos, Addison, TX

2002-2003, Administrative Assistant, Temporary (Various), Dallas, TX

2001, Service Express Attendant, Starwood Hotels, Westin Stonebriar, Frisco, TX

SKILLS

Microsoft Office Packages

Word

Powerpoint

Excel

Outlook

Programming

Python

PERL

UNIX

Igor Pro

Probability and Statistics

Laws of probability

Hypothesis testing

Propagation of error

Goodness-of-fit

Analysis of residuals

Maximum Likelihood estimation

Modeling and Mathematics

Formal logic

Recognizes the classic fallacies

ODEs/PDEs

Vector calculus

Derivations

Analysis and proof

Linear and nonlinear least-squares regression

General Laboratory Skills

ImageJ Image Processing

Researching Literature

Prioritization/Multitasking

Dilutions & Concentration Measurement

Buffers and Solutions

Molecular Biology

Cloning and Construct Design

PCR

Transformation

DNA extraction (alkaline lysis)

Restriction Digest

DNA gels

Microarray Analysis

Protein Science

SDS-PAGE

Inducible Expression

Cell Lysis/Centrifugation

Affinity Columns

Ion Exchange Columns

Size-Exclusion Columns

FPLC/HPLC

Practical Screening (Solubility, Folding, or Crystallization)

Fluorescent Labeling

Spectroscopic Methods

Analytical Ultracentrifugation (Beckman Optima XL-A)

Fluorimetry (ISS, Horiba)

UV-Vis

Microscopy

Confocal Microscopy

Scanning Electron Microscopy

PALM (rotation project)

Small-Angle Neutron Scattering

Contrast Variation Neutron Scattering

Mass Spectrometry

MALDI-TOF

LC/MS-MS

Gel band digestion and identification

Chemistry/Biochemistry

Basic organic synthesis

Reaction mechanisms

Folding assays

Bioinformatics

Sequence recovery and identification:

BLAST searches

NCBI resources

Organism databases

Uniprot

Global and local sequence alignment

Cladograms

Expasy resources

Sequence manipulation and analysis

Biophysics/Physics

Thermodynamic models

Kinetic models

Statistical mechanics

Physical forces

Molecular visualization

Familiar with biophysical methods

Teaching

Martial arts instructor

Tutor for organic chemistry

Teaching assistantships

Scientific Communication

One-on-one meetings

Lab meetings

Journal clubs

Seminars

Lectures

Posters

Publications

REFEREES

Professor Karen Fleming

Professor Bertrand Garcia-Moreno

Professor Richard Cone

RESEARCH

- Currently writing a thesis based on work on the structure, thermodynamics and kinetics of the model holding chaperone Seventeen Kilodalton Protein (Skp)
- Showed that Skp is not an obligate trimer, despite a pre-existing paradigm that Skp functions as a trimer
- Developed a simple but rigorous assay to describe holding activity kinetically using sedimentation velocity
- Interested in the science describing symbiosis of cells, and as a corollary the structure, biogenesis, and biological role of cell membranes
- Attended Delaware Membrane Symposium, Biophysical Society Conference, Institute for Biophysical Research retreat, Gordon Conference on Membrane Protein Folding, and Gibbs Conference on Biothermodynamics
- Presented several posters, and gave a talk at the Gibbs conference
- Contributed to three authorships, one of them a first authorship
- Developed an experimental description of the energetics and rates underlying Skp function in chaperoning outer membrane proteins

- Used sedimentation equilibrium and sedimentation velocity in parallel with intrinsic tryptophan fluorescence to quantify the Skp chaperone activity
- Utilized quantitative methodologies and modelling techniques backed up with statistics
- Developed skills in designing critical, practical experiments using appropriate methods to test specific hypotheses

TEACHING

- Teaching assistant, Computational Biology, for Professor Pat Fleming
- Teaching assistant, Cellular and Molecular Physiology, for Professor Richard Cone
- Participated in constructive critique of more than two dozen student second year seminar practice sessions
- Provided advice and mentorship for first and second year graduate students

ADMINISTRATION

- Session chair, Molecular Chaperones, Gordon Conference on Membrane Protein Folding
- Graded homework
- Graded and proctored exams
- Provided feedback on publication and grant proposal drafts
- Practiced prioritization and multitasking
- Gained skills with project feasibility assessment (when to keep going and when to shift focus)
- Maintained instrumentation and other hardware while working closely with service technicians
- Purchased laboratory hardware and worked with vendors and salespeople

PUBLICATIONS

Zaccai NR, Sandlin CW, Hoopes JT, Curtis JE, Fleming PJ, Fleming KG, Krueger, SK. Deuterium Labeling Together with Contrast Variation Small-Angle Neutron Scattering Suggests How Skp Captures and Releases Unfolded Outer Membrane Proteins. *Methods in Enzymology.* **566:159-210** (2016)

Sandlin, CW, Fleming, KG,.Skp trimer formation is insensitive to salts in the physiological range. *Biochemistry.* **54(48):7059-62** (2015)

Gessmann D, Chung YH, Danoff EJ, Plummer AM, Sandlin CW, Zaccai NR, Fleming KG. Outer membrane β-barrel protein folding is physically controlled by periplasmic lipid head groups and BamA. *Proc Natl Acad Sci.* **111(16):5878-83** (2014)

ORAL PRESENTATIONS

Sandlin CW, Skp trimer formation is insensitive to ionic strength. 29th Gibbs Conference on Biothermodynamics (2015)

POSTER PRESENTATIONS

Sandlin CW, Zaccai NR, and Fleming KG. Skp trimer formation is insensitive to ionic strength. 1st Gordon Conference on Membrane Protein Folding (2015)

Sandlin CW, Zaccai NR, and Fleming KG. Skp revisited: Trimerization is thermodynamically linked to binding and is salt-independent. 28th Gibbs Conference on Biothermodynamics (2014)

Sandlin CW and Fleming KG. *E. coli* chaperone Skp has diverse effects on membrane protein folding efficiency in vitro, and displays self-association that is robust to salt. Institute for Biophysical Research Retreat (2013)

CONFERENCES

2011-2015, Institute for Biophysical Research Retreat

2013-2015, Delaware Membrane Symposium

29th Gibbs Conference on Biothermodynamics

1st Gordon Conference on Membrane Protein Folding

59th Biophysical Society Conference

28th Gibbs Conference on Biothermodynamics

PROFESSIONAL MEMBERSHIPS

Biophysical Society

The Honor Society of Phi Kappa Phi